Collective Behaviour of Crimes In Cyberspace

Andrzej Buda
Interdisciplinary Research Institute, Głogów, Poland
ul. Oriona 15/8, 67-200 Głogów, Poland
E-mail address: andrzejbuda@wp.pl
ORCID: https://orcid.org/0000-0002-2492-5580

Abstract

Aim. The Convention of Cybercrimes (ETS No. 185) singned in Budapest, Hungary in 2001 created the unification of cyber-crimes and had impact on national criminal codes across the world. Therefore I started to colaborate with the National Police Headquarters from Warsaw to investigate direct and indirect connections between cyber-crimes defined by the Convention and the Criminal Code in Poland because statistical analysis of cyber-crimes were not thus far satisfactory .

Methods. Statistical analysis of data collected in time series may lead to computer correlation coefficients between crimes and detect a hierarchical structure of selected cyber-crimes

Results. According to correlation coefficients between crimes, the behaviour of independent cyber criminals is collective and might lead to self-organised criticality. Correlations and anticorrelations between crimes are extremely strong. Therefore crimes may exist collectively or may exclude each other

Conclusion. Hierarchical structure of crimes according to the MST seems to be extremely logical and might lead to a perfect crime plans or prevention against cyber-crimes as well. Therefore the sensitive details given by graph of the MST had to be censored.

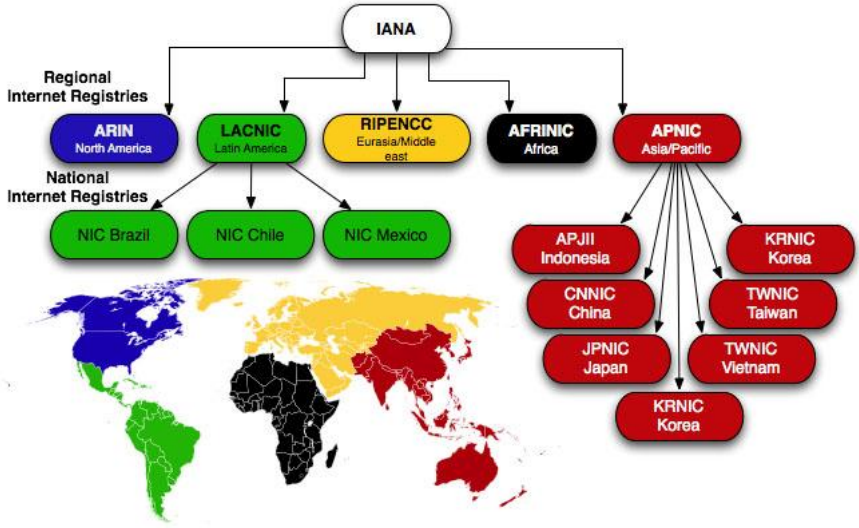
Keywords: Cyber-crimes, correlations, anticorrelations, time series analysis, Minimum Spanning Tree, Socual networks, Criminal law

Introduction and motivation

The global nature of the Internet determines a diffusuon of information (Di Sia, 2015, Kappelar 2010). On November the 23rd 2001 the Convention of Cybercrimes (ETS No. 185) was singned in Budapest, Hungary. The Convention was the first international treaty on crimes committed via the Internet and other computer networks, dealing particularly with infringements of copyright, computer-related fraud, child

pornography and violations of network security. It also contains a series of powers and procedures such as the search of computer networks and interception. The main aim of creating this Convention was the global nature of crimes that had to be defined by unification because the application of national criminal codes was not efficient.

Fig. 1. National structure of IP adressess



Computer crime encompasses activities, including **computer fraud**, **financial crimes**, **scams**, **cybersex trafficking**, and ad **fraud**. For example, the keylogger - a spyware that can covertly record your keystrokes – might be used nowadays by tracking and overhearing the frequencies of sounds created by pressing digits. The data captured can be sensitive information that is used to commit fraud or identity. Forms of computer fraud include **hacking** into computers to alter information, **distributing malicious code** such as computer worms or viruses, installing **malware** or **spyware** to steal data, **phishing**, and **advance-fee scams**. Other forms may be committed using computer systems (bank fraud [Schmitt 1999, Mizuno 2003], carding, identity theft, extortion, and theft of classified information). Methods like **hacking** and **sniffing** are also classified as cyber-crimes. [Stapór 2020].

The new sophisticated methods of cyber-crimes, however, did not change the core of the Criminal Code (**kk**) in Poland [Makowski 1937, Piórkowska-Flieger 2004, Lelental 1977]. However, some additional articles had to appear as an extension of regulations that already existed. Therefore I could collect monthly data for all the cyber-crimes commited in Poland between Jan. 2016 and Dec. 2020.

Cyber-crimes according to Criminal Code in Poland (2016-2020)

In my research, I collected database of time series and computed main statistical features:

Table 1.
Statistical properties of cyber-crimes in Poland (2016-2020)

Article (Criminal Code)	Mean	Mediana	Standard deviation	Min.	Max
267 § 1 kk	2910	2008	1368	1834	4595
267 § 2 kk	171	170	55	113	235
267 § 3 kk	48,6	47	9,71	40	65
267 § 4 kk	7,8	7	5,97	3	18
268 § 1 kk	71,8	72	7,40	64	81
268 § 2 kk	28,2	28	9,36	17	41
268a § 1 kk	446	430	120	298	616
268a § 2 kk	2	2	1,87	0	5
269 § 1 kk	5,4	6	2,07	2	7
269a kk	30,6	30	7,50	19	39
269b kk	34,4	37	7,23	24	41
286 § 1 kk	5869	5083	2712	3295	9088
286§1 w zw.z 294§1kk	11,8	13	2,95	7	14
286 § 2 kk	35,2	28	19,6	17	60
115 ust. 1 ustawy p.a.	11,6	10	4,39	8	19
115 ust. 2 ustawy p.a.	4,6	5	0,548	4	5
115 ust. 3 ustawy p.a.	4,6	4	2,41	2	8
116 ust. 1 ustawy p.a.	2233	2384	784,6	1186	3203
116 ust. 2 ustawy p.a.	2306	912	4007	24	9417
116 ust. 3 ustawy p.a.	2924	1174	3238	67	7879
116 ust. 4 ustawy p.a.	726	9	1111	1	2529
117 ust. 1 ustawy p.a.	177	59	220	8	503
117 ust. 2 ustawy p.a.	0,6	0	0,894	0	2
118 ust. 1 ustawy p.a.	47,2	19	77,3	6	185
118 ust. 2 ustawy p.a.	2,6	1	4,72	0	11
118 ust. 3 ustawy p.a.	11,0	1	16,3	0	38
118(1)ust.1ustawy p.a.	2,8	3	2,68	0	7
118(1)ust.2ustawy p.a.	0,8	1	0,837	0	2
119 ustawy p.a.	0,2	0	0,447	0	1
200 § 1 kk	37,8	40	6,50	30	45
200 § 3 kk	225	234	69,2	128	298
200 § 4 kk	22	21	8,75	12	34
200 § 5 kk	0,600	1	0,548	0	1
200a § 1 kk	76,2	71	17,3	61	106
200a § 2 kk	373	373	96,9	245	503
200b kk	15,4	4	27,8	0	65
202 § 1 kk	56,8	53	25,8	21	91
202 § 3 kk	1396	1243	1027	443	3040
202 § 4 kk	3511	232	5345	180	12490
202 § 4a kk	23,6	20	9,71	15	40
202 § 4b kk	26,0	17	23,3	8	66
202 § 4c kk	0,400	0	0,548	0	1

According to this data, the most frequent crime is hacking. On the other hand, thwarting access to copyright information does not occur at all. In 2020 the number of cyber-crimes in Poland increased. The Covid-19 public restrictions might be an explanation for this phenomenon because each human traditional activity was limited and the Internet was the last field free from boundaries.

In my research, I created the Correlation matrix *34x 34* for all the crimes, where the correlation coefficient was defined as:

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y}$$

where cov*(X,Y)* refers to variables *X* i *Y*, and *σx* or *σy* are standard deviations of these variables. The total number of 1156 matrix elements may be reduced to 536 coefficients because this triangle matrix is symmetric and all the diagonal elements are equal to 1.

According to this, the strongest detected correlations [Buda, 2010] lead to conclusions:

- Creating devices for breaking passwords is strongly correlated with causing a huge damage in data base systems (0.99)
- Benefits from digital piracy become constant income source (0.98)

On the other hand, the strongest anticorrelations were detected as well. In consequence:

- Daring misleading excludes hiding information on copyrights (-0.91) and sniffing (-0.90)
- Hacking excludes damage in the system (-0.89)

It is also possible to create the Minimum Spanning Tree for the most popular cyber crimes. The Minimum Spanning Tree method based on correlation coefficients has been useful in economics and physics since 1999 [Mantegna, 1999]. This tool, however, has been overlooked by other data miners. The purpose is to provide a detailed method of detecting hierarchical structure of the given data sets including direct and indirect connections between elements. It may be applied by:

1. Choosing the collected dataset and time-series
2. Computing the matrix of correlations between *n* elements
3. Creating Minimum Spanning Tree that has *n* nodes and *n*-1 connections by Kruskal or Prim's algorithm [Kruskal, 1956]

The Minimum Spanning Tree may reveal the structural properties of cyber-crimes.

Firstly, I had to detect the pair with the strongest correlation coefficient. In my research, the strongest pair (0.99) is between creating devices that may break protected systems (268a § 2 kk) and the digital piracy (Art. 118(1) ust. 1 ustawy p.a.). This connection might be described by crimilal codes articles as:

268a § 2 kk - Art. 118 (1) ust. 1 ustawy p.a.

Another strongest coefficient is the one between the digital piracy and the income from digital piracy (Art. 118 ust. 2 ustawy p.a.) with 0.98. Therefore the first strongest connections may be plotted as a chain:

268a § 2 kk - Art. 118 (1) ust. 1 ustawy p.a.-Art. 118 ust. 2 ustawy p.a.

The complete (censored) Minimum Spanning Tree obtained for 34 selected cyber-crimes is given in Fig 2. Hierarchical structure of cyber-crimes is valuable for the National Police Headquarters in Poland to prevent commiting another crimes because of direct and indirect ifluences between crimes in their nature. On the other hand, the original MST is extremely logical and might lead to potential perfect crimes. Therefore I decided to publish the censored MST only.

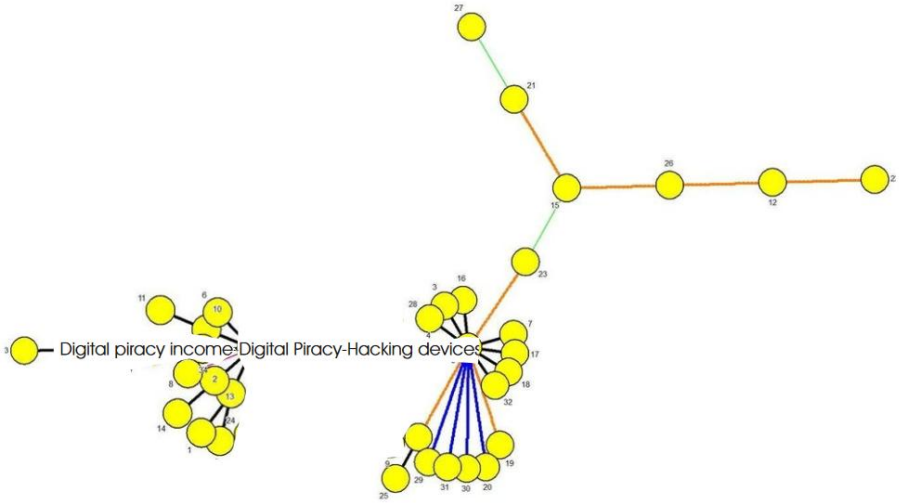


Fig. 2. Hierarchical structure of cyber-crimes in the Minimum Spanning Tree

Another dangerous feature of cyber-crimes is a collective behaviour of thousands independent criminals, because they act like a flock of seagulls. Such a phenomena is well known in complex systems of physics, finance or biology and described as self-organised criticality [Bak, 1987]. It occurs because:

- the majority of correlations between crimes revealed by the Minimum Spanning Tree is extremely strong
- criminals that commit cyber-crimes all over Poland are usually independent

Conclusion

Statistical analysis of cyber-crimes from a correlation coefficients' point of view leads to dangerous conclusions because independent criminals may commit crimes that are extremely logical. Therefore the complete uncensored MST may look like a plan for commiting a perfect crime. However, the hierarchical structure of cyber-crimes may be useful for policemen because it is easier to prevent

by using knowledge about direct or indirect influences between cyber-crimes on each other.

On the other hand, the existence of extremely strong anticorrelations is also valuable because it provides information about contradictions between crimes. After finding collective behaviour, I would later propose models of ferromagnetism with external field because using the Ising models [Ising, 1925] may seem more reasonable in this area.

Acknowledgement

Author would like to thank the National Police Headquarters (Poland) for collecting and providing data.

References

Bak P., Tang C., Wiesenfeld K. (1987) Self-organized criticality: an explanation of 1/f noise *Physical Review Letters* 59 (4), 381–384.

Buda, A., Jarynowski, A. (2010). *Life-time of correlations and its applications vol. 1.* Wrocław: Wydawnictwo Niezalezne.

Di Sia, P. (2015). About the Internet and the diffusion of science, *E-methodology, 2,* 18-26; doi: 10.15503/emet2015.18.26

Ising E (1925), Contribution to the Theory of Ferromagnetism, *Zeitschrift fur Physik, vol. XXXI*, 253-258

Kappeler, P. M., Silk J. B. (2010) *Mind the Gap: Tracing the Origins of Human Universal, Springer, Heidelberg,* 19-51

Kruskal J. B. (1956), On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem, *Proceedings of the American Mathematical Society, 7*, (1), 48–50

Kwapień, J., Drożdż, S. (2012). Physical approach to complex systems. *Physics Reports Volume, 515*(3–4), 115-226.

Lelental S, Górniok O., Popławski H. (1977) *Prawo karne. Część szczególna* [Special Cases In Criminal Code], Gdańsk: Uniwersytet Gdański.

Makowski W (1937), *Kodeks karny. Komentarz* [Criminal Code. Introduction], Warszawa: Monolit.

Mantegna R. N. (1999), Hierarchical Structure In Financial Markets, *Condensed Matter and Complex Systems, 11,* 193-197.

Mizuno T., Kurihara S., Takayasu M., Takayasu H.(2003), Analysis of high resolution foreign exchange data of USD-JPY for 13 years. *Physica A: Statistical Mechanics and its Applications 324 (1),* 296-302.

Piórkowska – Flieger J. (2004), *Fałsz dokumentu w polskim prawie karnym* [Counterfeit According To Polish Criminal Law], Kraków: KUL.

Schmitt, F., Schertzer, D., & Lovejoy, S. (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models and Data Analysis, 15(*1), 29-53.

Stąpór A. (2020) Europejski miesiąc cyberbezpieczeństwa [European Month Of Cyber-Security], *Policja 997*, 10 (87), 4-7.

West D. B. (1996), *Introduction To Graph Theory*, Prentice-Hall, Englewood Cliffs NJ