

Wavelet and fractal analysis based news spreading model

B. L. Smetanyuk¹, V. V. Mishyn¹, Y. V. Nakonechna¹

¹*National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»,
Institute of Physics and Technology*

Abstract

The aggression of Russian Federation in Ukraine has exposed a strategic problem, derived from unwillingness to fully oppose in the information confrontation. Media could do even more damage to country's defense capabilities than military action. Spreading fake news can be used to misrepresent, distort, and create a negative image of an informational event in society, as well as instill fear, panic, or downplay achievements. Little attention is paid to such information operations and approaches to their detection. The paper deals with detection of such information effects using Exploratory data analysis, fractal and wavelet analysis, as well as factor structure analysis using neural networks for being defensive, understanding of methods and techniques used to attack in Informational Space is necessary.

Keywords: Fake news detection, fractal analysis, wavelet, EDA, ML, cybersecurity

Problem formulation

Literary sources analysis. At the moment, «*fake news spreading*» is actively discussed by a professional community of researchers as definite and meaningful action. Fake is set of fictitious and specially fabricated news whose main purpose is to undermine the reputation of targeted institution, organization or person. Such news are destructive, causing financial damage and potentially harmful to state, citizens, companies, organizations.

In the context of ongoing «*hybrid war*» it is concluded, that systemic misinformation and manipulation of audience poses a threat to Ukrainian national interests and security and also and discredits media outlets. As such, for fake news information protection requirements are violated, making them indirectly malicious. Their influence spread peaked with mainstream embracing use of social media and subsequent automatic distributing by different organizations. Such an example was botnet generating and posting misinformation about US being responsible for COVID-19 creation and spread.

Fake news detection methods can be divided in two categories, according to source of their distribution: the media news content research and social networks content research. Methods related to the first category are focused on the news content: text, header and additional metadata (when they are available). The second category methods are focused on social networks features (user interaction analysis and user's activity monitoring). It is possible to distinguish different approaches to news research and fake news management in mentioned research categories: fake news monitoring, linguistic cue methods and text analysis, network analysis, analytic approaches including machine learning methods.

The news content approach describe features connected with the meta information related to a piece of news. Representative news content attributes are

source, headline, body text, photo/video sources. The main tool for news content analysis are fact-checking and text mining techniques.

The social network content approach. Generally, there are three major aspects of the social media context represented: users, generated posts, and networks. Social engagements represent the news proliferation process over time, which provides useful auxiliary information to infer the veracity of news articles based on graphic representation of users connections, news spreading model and news spreading characteristics, users activity on post news etc [1, 2].

Main material presentation

Goals and methods of studying

The subject of the study is information attacks and information influences. The aim is to develop an approach to identify information effects based on a model built using wavelet and fractal analysis for further implementation in specialized systems of intellectual analysis of unstructured text flows on the Internet. This would give an opportunity to defense experts or law enforcement to detect spikes in the information field and the dissemination of manipulative messages and allow to react in a timely and effective manner. To perform the study on informational influences, the informational messages set on the news related to the granting the Tomos of autocephaly to Ukraine. Collected data was analyzed step by step with following methods: Exploratory Data Analysis (using statistical processing on time series composed from number of published news sorted by emotional affiliation) [3, p.20], Fractal analysis [4, p.112], [3, p.154], Wavelet analysis [4, p.115], [3, p.179], Factor structure analysis [5, 6].

Step 1: Exploratory Data Analysis

Exploratory data analysis (EDA) provides patterns revelation and dataset's main characteristics summarizing beyond modeling and hypothesis testing. Results after EDA performing are outliers and anomalies detection, underlying structures uncovering, statistically analyzed informational time series. Thus, exploratory data analysis is used as a preliminary step before the hypothesis testing stage. In this case the main EDA purpose is to identify the characteristics, patterns, and relationships in the data that are the subject of the analysis [1, 2].

As time series analysis subject was used the number of news about Tomos. The set of analyzed sources for data collection consisted of 20 television channels, 400 traditional media, 14000 online media, 15 radio stations, and social networks presented by Facebook, VKontakte, YouTube, Odnoklassniki, Live journal. The resulting time series characterize the total amount of news, classified by different emotional color, from different sources from the Ukrainian and Russian media, and also as well as separate regions of Donetsk and Lugansk regions (SRDLR) media. Emotional color classification was performed using Sentiment analysis (also known as opinion mining or emotion AI) which refers to the use of natural language processing and text analysis to systematically identify and quantify affective states and subjective information.

After Exploratory data analysis performed on classified data following time series data characteristics were identified:

- Mean
- Median
- Mode
- Range
- The first quartile or 25th percentile, which corresponds to the value that lies halfway between the median and the lowest value in the distribution (when it is already sorted in the ascending order)
- The third quartile or 75th percentile, which corresponds to the value that lies halfway between the median and the highest value in the distribution (when it is already sorted in the ascending order)
- Inter-Quartile Range
- Dispersion
- Standard deviation
- Coefficient of variation

The mean value makes it possible to trace the fluctuations of the time series values, and provides indirect information on the impact of data emission knowing the Mode value.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \tag{1}$$

where μ is the mean, x_i is time series item value.

The outliers are values in data that are abnormal. In case, abnormal nature is considered that the values are significantly different from others. Hypothetically, news time series outliers characterize the interference of

fake news in the social processes of dissemination and exchange of information.

Inter-Quartile Range is the difference between the extreme values of the sample, truncated on the 25th to the 75th quartile. Such a measure is less prone to the impact of emissions and potentially characterizes the real amplitude of fluctuations in the social process of dissemination and exchange of information.

Dispersion and standard deviation characterize the scattering of values around the data center.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \tag{2}$$

where s^2 is dispersion, $(x_i - \bar{x})$ is deviation from mean value.

Standard deviation is the values scattering index.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \tag{3}$$

The coefficient of variation gives an idea of the size of the deviation of the sample values from each other and allows to compare the variance in the values between different points in the data.

The result of time series analysis on news set data is represented in Tabs 1, 2, 3:

Table 1. Statistical characteristics of Ukrainian sourced news

	All messages	Positive	Negative	Neutral
Mean	333.2	233.4	93.26	6.53
Median	178	118	54	2
Mode	0	0	0	0
Range	4570	3743	976	250
25th percentile	72	46	20	0
75th percentile	407	264	119	7
Inter-Quartile Range	335	218	99	7
Dispersion	245030.4	143628.8	14359.3	292.4
Standard deviation	495	378.9	119.8	17.1
Coefficient of variation	73539.7	61537.6	15397.1	4474.9
Total messages	188255	131871	52692	3692
Percent of all	100	70.04	27.98	1.96
Max value a day	4570	3743	976	250

Table 2. Statistical characteristics of Russian sourced news

	All messages	Positive	Negative	Neutral
Mean	69.3	36.6	31.4	1.19
Median	33	17	16	0
Mode	0	0	0	0
Range	841	404	464	19
25th percentile	14	7	5	0
75th percentile	83	43	38	1
Inter-Quartile Range	69	36	33	1
Dispersion	9341.8	2904.14	2039.8	5.42
Standard deviation	96.6	53.9	45.16	2.33
Coefficient of variation	13487.9	7928.29	6488.131	455.4
Total messages	39132	20696	17763	673
Percent of all	100	52.9	45.4	1.7
Max value a day	841	404	464	19

Table 3. Statistical characteristics of SRDLR sourced news

	All messages	Positive	Negative	Neutral
Mean	6.7	2.8	3.8	0.1
Median	4	2	2	0
Mode	0	0	0	0
Range	97	40	56	6
25th percentile	1	0	1	0
75th percentile	6	4	5	0
Inter-Quartile Range	8	4	4	0
Dispersion	84	16.4	29.5	0.2
Standard deviation	9.2	4	5.4	0.5
Coefficient of variation	1243.2	589.1	762.6	184.2
Total messages	4125	1695	2357	73
Percent of all	100	41.1	57.1	1.8
Max value a day	97	40	56	6

As can be seen from Tables 1-3, the most popular value that characterizes the number of news per day is 0. This is evidenced by the Mode indicator of these time series. It can also be seen that the Inter-Quartile Range is approximately equal to the average number of messages per day. This makes it possible to make more accurate predictions and hypotheses around outliers source. Using mentioned results it can be seen that SRDLR news is similar in structure to Russian news due to time series characteristics and the ratio of news with a negative color, which much higher than in Ukrainian news. In both Russian and SRDLR news time series, the ratio of news with a negative color is much higher than in Ukrainian news, and also with a greater concentration of positive news, the maximum number of negative news per day is higher than positive.

Graphic representation of performed analysis of emotional distribution are below:

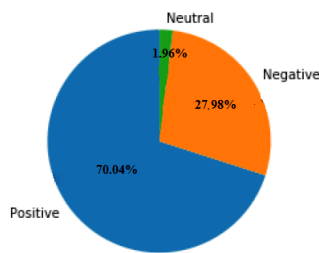


Fig. 1. Round diagram of news distribution by emotion in Ukrainian sources

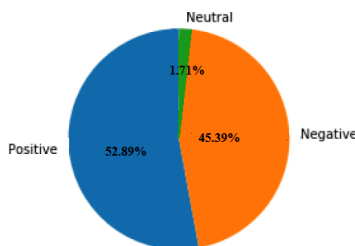


Fig. 2. Round diagram of news distribution by emotion in Russian sources

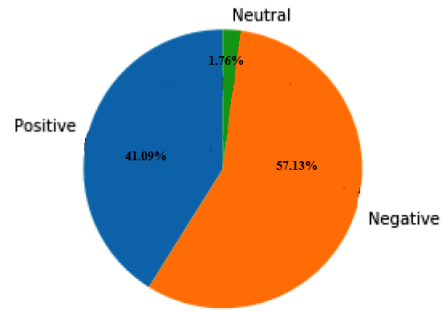


Fig. 3. Round diagram of news distribution by emotion in SRDLR sources

Statistical characteristics of the considered time series are presented for each source by box plot as a method for graphically depicting groups of numerical data through their quartiles:

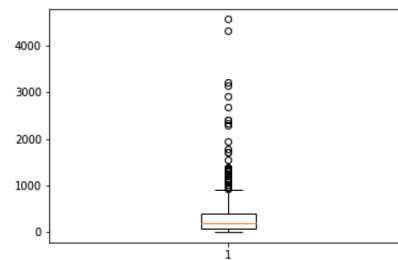


Fig. 4. Box plot describing news posted by Ukrainian sources

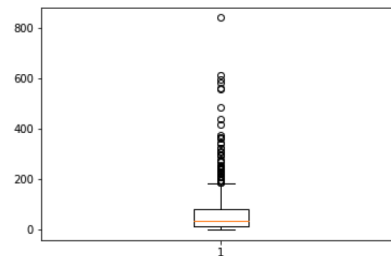


Fig. 5. Box plot describing news posted by Russian sources

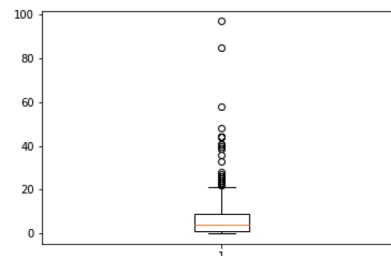


Fig. 6. Box plot describing news posted by SRDLR sources

A general view of the obtained time series characterizing amount of publications with different sentiment score per day between March 2018 and October 2019 is presented in the Figures 7, 8, 9:

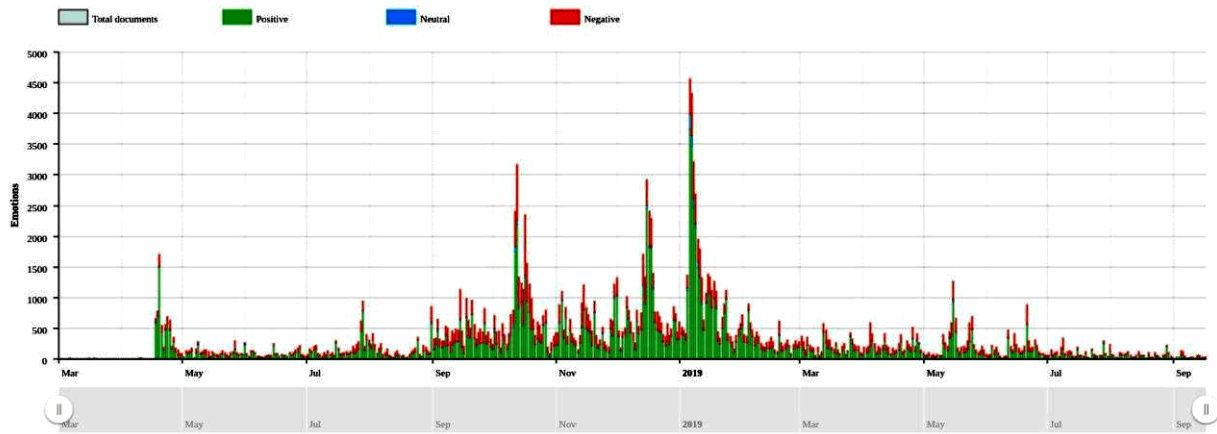


Fig. 7. Statistical distribution of classified by emotion Ukrainian news

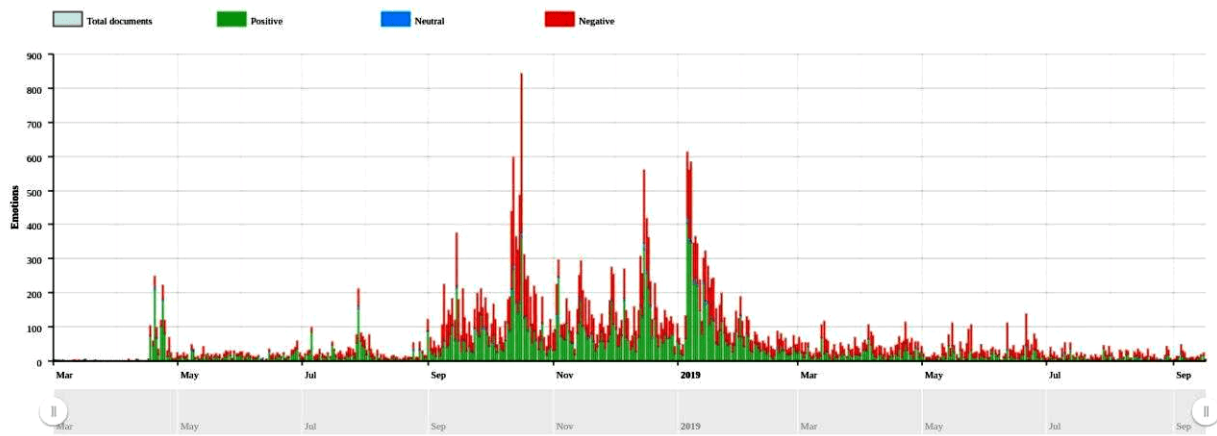


Fig. 8. Statistical distribution of classified by emotion Russian news

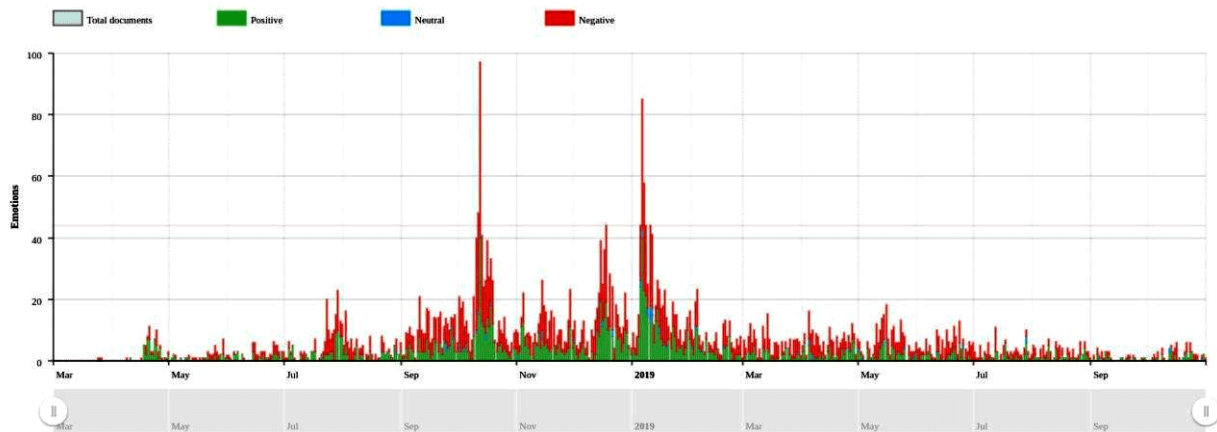


Fig. 9. Statistical distribution of classified by emotion SRDLR news

The box plot shows that for all three regions there are anomalous outliers related to certain dates. There are two main outlier points for Ukrainian news corresponding to the January 4, 2019 and January 5, 2019. These dates are characterized by an abnormally high number of positive news. The largest recorded values for positive news per day of all time are probably related to Tomos signing announce and successful Tomos signing. For Russian news, such a point is one and

it corresponds to October 15, 2018. This day is also characterized by an abnormally high amount of positive news, but also a number of negative news was recorded. The SRDLR outliers are corresponding to the dates October 11, 2018 (the Ecumenical Patriarch Bartholomew in Constantinople confirmed that Ukraine will be granted the Tomos – meaning that Ukraine’s independent autocephalous church is recognized) and January 5, 2019. Both of these days are also character-

ized by an increase in positive news, and their number fixed as maximum. In addition, the largest amount of negative news was recorded on October 11.

Step 2: Fractal Analysis

As the information space is considered as stochastic, the self-similarity of the information space is expressed in stable structure in which such sections as sources, authors, subjects practically do not change the form. Fractal theory application in the informational space analysis allows to investigate the patterns that form the basis of computer science. Due to thematic information arrays are represented by information clusters, it is possible to perform a cluster analysis and independently identify new features objects and develop information space self-similar structures as stochastic fractals distribute objects to new groups [3, p.158].

Within the applied fractal analysis, following indicators were calculated:

- the Hurst exponent;
- Fractal dimension;
- Correlation entropy;
- Correlation dimension.

Hurst exponent H is a persistence measure characterizing a process property to trending. The value $H > \frac{1}{2}$ means that a certain side dynamics of the process in the past is likely to cause future dynamic direction. In $H < \frac{1}{2}$ case it is predicted that the process will change a direction. $H = \frac{1}{2}$ means process uncertainty [3, p.164]. The Hurst exponent is also known as R/S analysis. To study the fractal characteristics of time series the values of the Hurst exponent are investigated from the ratio:

$$(R/S) = (A \times N)^H$$

where

- R is the range between the minimum and maximum value;
- S is the standard deviation;
- A is the scale constant;
- N is the sample size;
- H is the Hurst exponent.

Thus, Hurst exponent is calculated as:

$$H = \log_{(R/S)} (A \times N)$$

Fractal dimension determines the degree of complexity of the fractal figure and described by ratio

$$D = \frac{\log n}{\log r}$$

where n is the number of self-similar parts after figure is r -times increased.

The correlation entropy index is required to analyze the dynamics of the process subject to predictability. The smaller the value of the correlation entropy, the more predictable the behavior of the process. In other words, correlation entropy is a quantitative indicator of the chaos of the social process. The correlation dimension indicator shows complexity of the process behavior.

Thus, the correlation dimension indicates presence of similar values in the cluster system, probability that a random number of the time series will spend in this area, and also shows the number of parameters required process describing [6].

The result of fractal analysis on news set data is represented below:

	All messages	Positive	Negative	Neutral
Hearst index	1.0319 ± 0.1473	1.0160 ± 0.1651	1.0198 ± 0.1233	0.7308 ± 0.1753
Fractal dimension	0.9681 ± 0.1473	0.9840 ± 0.1651	0.9802 ± 0.1233	1.2692 ± 0.1753
Correlation entropy	0.422	0.219	1.080	0.254
Correlation dimension	1.973	2.654	4.271	0.227

Fig. 10. Time series indicators for Ukrainian news

	All messages	Positive	Negative	Neutral
Hearst index	1.1025 ± 0.1815	1.0882 ± 0.2125	1.0381 ± 0.1538	0.8616 ± 0.3143
Fractal dimension	0.8975 ± 0.1815	0.9118 ± 0.2125	0.9619 ± 0.1538	1.1984 ± 0.3143
Correlation entropy	0.450	0.091	0.436	0.084
Correlation dimension	2.935	4.393	4.627	0.000

Fig. 11. Time series indicators for Russian news

	All messages	Positive	Negative	Neutral
Hearst index	0.8605 ± 0.1579	0.8893 ± 0.2365	0.8706 ± 0.1932	0.6624 ± 0.1973
Fractal dimension	1.1395 ± 0.1579	1.1107 ± 0.2365	1.1294 ± 0.1932	1.3376 ± 0.1973
Correlation entropy	0.024	0.102	0.228	0.000
Correlation dimension	0.412	0.322	0.000	0.000

Fig. 12. Time series indicators for SRDLR news

Presented results show that all processes are persistent, except the neutral SRDLR news, which trend is difficult to predict. As for Ukrainian news, the most uncertain is the process of news with a negative emotional color. The value of correlation entropy for positive and negative news is approximately equal. For Russian news, the result is the same, with the values of correlation entropy are much smaller. For SRDLR news, the correlation entropy for neutral news is 0. This means that the process that characterizes neutral news is completely defined and predictable.

Graphic interpretation of gained results is given below (Ukrainian news on Fig. 13, Fig. 14, Russian news on Fig. 15, Fig. 16, SRDLR news on Fig. 17, Fig. 18):

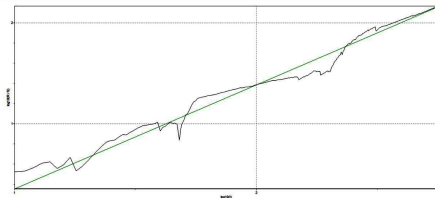


Fig. 13. Hurst exponent and Fractal dimension. Ukrainian news

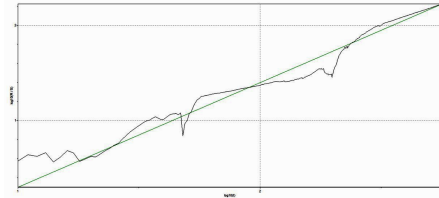


Fig. 15. Hurst exponent and Fractal dimension. Russian news

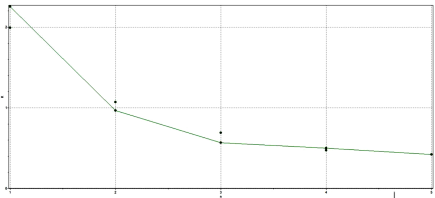


Fig. 14. Phase space dimension. Ukrainian news

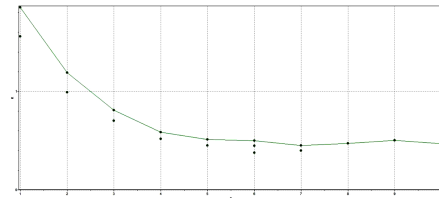


Fig. 16. Phase space dimension. Russian news

Step 3: Wavelet Analysis

The wavelet transform's main idea is non-stationary time series division into separate intervals (so-called observation windows) and performing scalar calculation on each of them gaining a value that characterizes two data patterns closeness degree, with different wavelet shifts on different scales. The wavelet transform generates a set of coefficients as functions of two variables: time and frequency, and therefore form a surface in three-dimensional space. These coefficients show to what extent the behavior of the process at a given point is analogous to wavelet on this scale. Mentioned operations use allows to analyze data at different scales and accurately mean the position of their characteristics over time [4, p.113]. The technology of using wavelets allows to detect single and non-regular «bursts», sharp changes in quantitative indicators values, thematic publications amount on the Internet in particular case.

News posting process can be described as oscillations, which are in some sense regular. Thus, fluctuation amplitude will indicate published at a given point in time news amount, taking into consideration specific news characteristics as first rapid increase, and then a slower decline. Since at rest the fluctuations of the social process are regular, interference in this process of fake news creates noticeable jumps and disrupts the structure of fluctuations. A study is also possible in which news that does not have a fake nature is considered noise. In this case, it is possible to extract the fake signal from the general news stream by separating the signal from the noise. After fractal analysis the signal was investigated for the property of self-similarity, wavelet analysis shows time series cycles inheritance and investigate these cycles for the moment of their beginning to the end [3, p.179], [4, p.113].

In study were used first- and fifth-order Dobesi wavelets, which differ in filter lengths of 2 and 10. Dobesi Wavelet is an orthogonal wavelets family with a compact carrier that is calculated iteratively. The main purpose of Dobesi wavelet use is to highlight media data patterns in Russian, Ukrainian and SRDLS news.

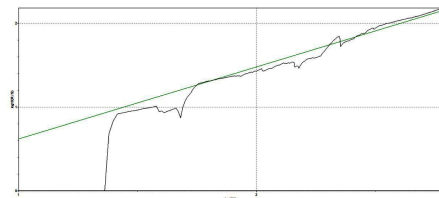


Fig. 17. Hurst exponent and Fractal dimension. SRDLR news

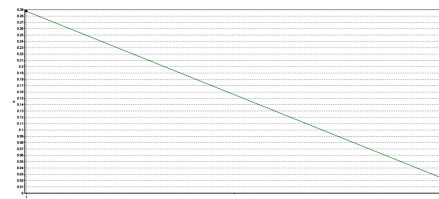


Fig. 18. Phase space dimension. SRDLR news

The task is to select similar cycles in all time series. Such a result has a very useful effect on the further construction of a model of fake news, as well as allow to refute or confirm some hypotheses [7].

The following result is a 10-level time series data decomposition by first- and fifth-order Dobesi wavelets. The 10th decomposition level is the maximum possible series representation level. The 11th level completely repeats the 1st. All levels decomposition signals are presented in details of the time series form and its approximate form. Graphic interpretation of gained results is given below (Ukrainian news on Fig. 19, Fig. 20, Russian news on Fig. 21, Fig. 22, SRDLR news Fig. 23, Fig. 24):

Due to gained time series decompositions it can be seen that Ukrainian news almost completely repeats the dynamics of positive news at each decomposition level (as positive news amount is about 70 percent). This means that the sharp changes in negative news amount are atypical for these time series and can be considered as information influence result.

Russian news general dynamics is a composition of positive and negative news. The situation directly cor-

Wavelet and fractal analysis based news spreading model

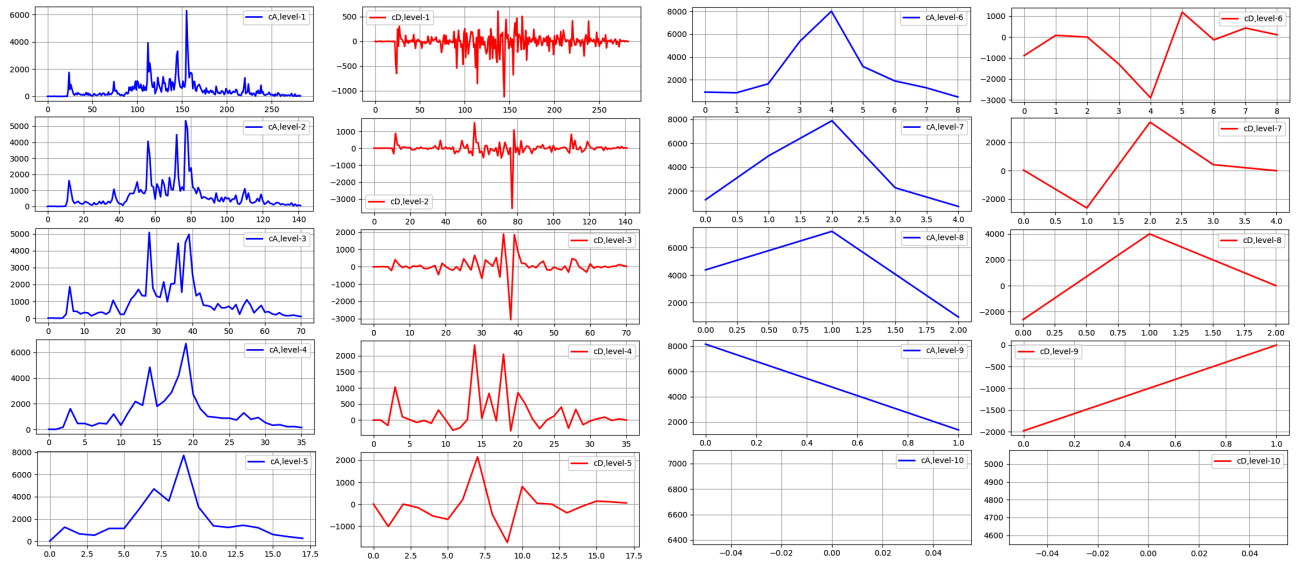


Fig. 19. First-order Dobesi wavelet decomposition (1-10 levels). Ukrainian news

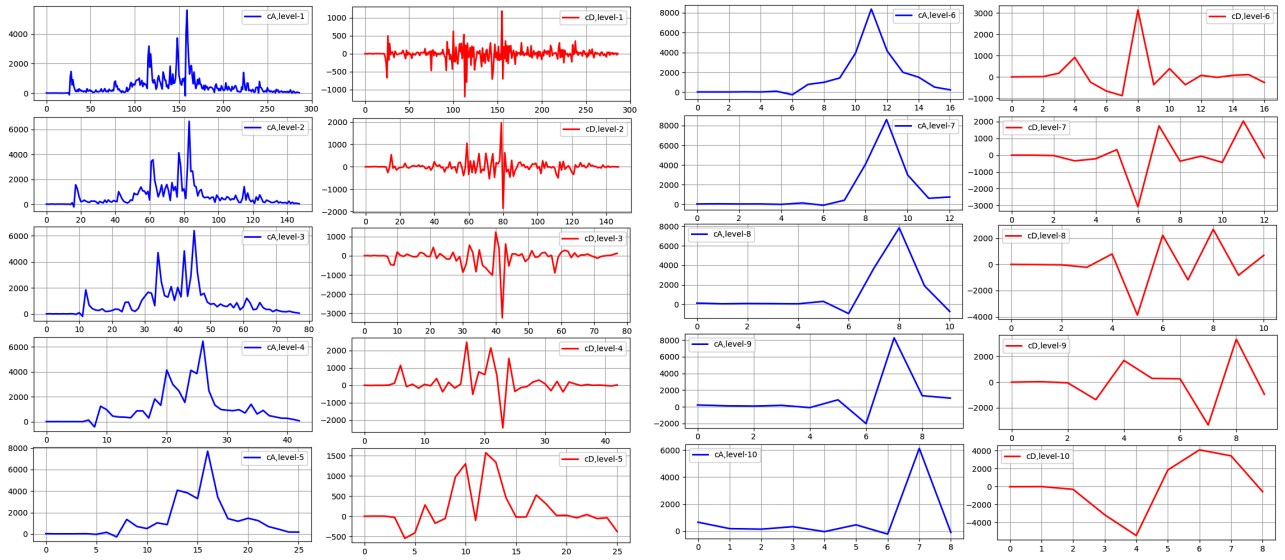


Fig. 20. Fifth-order Dobesi wavelet decomposition (1-10 levels). Ukrainian news

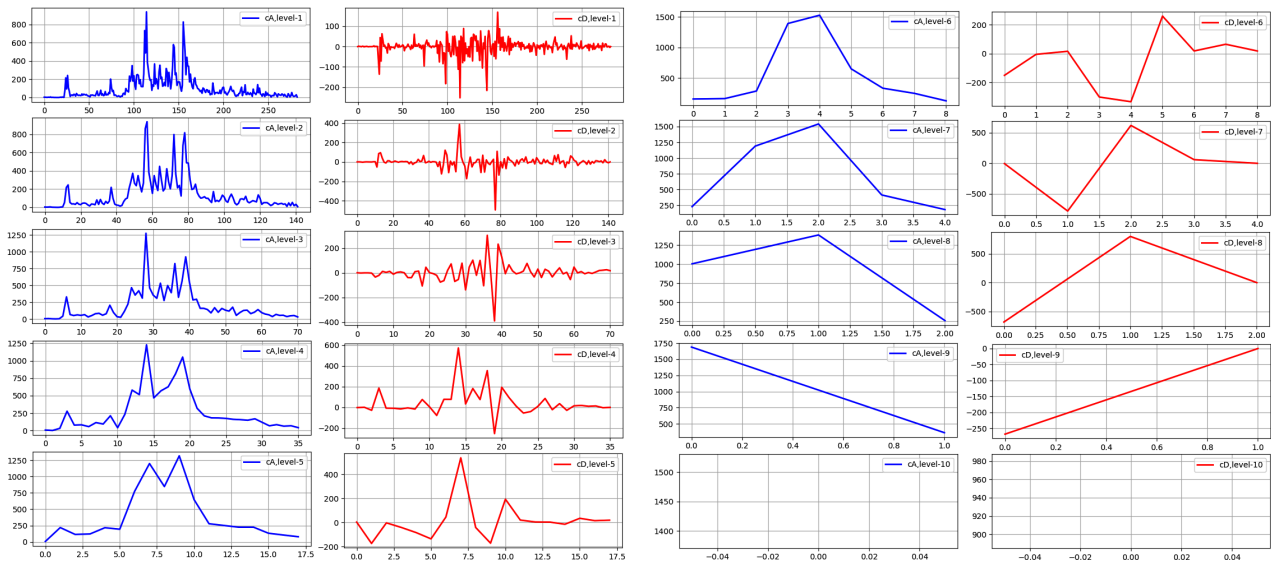


Fig. 21. First-order Dobesi wavelet decomposition (1-10 levels). Russian news

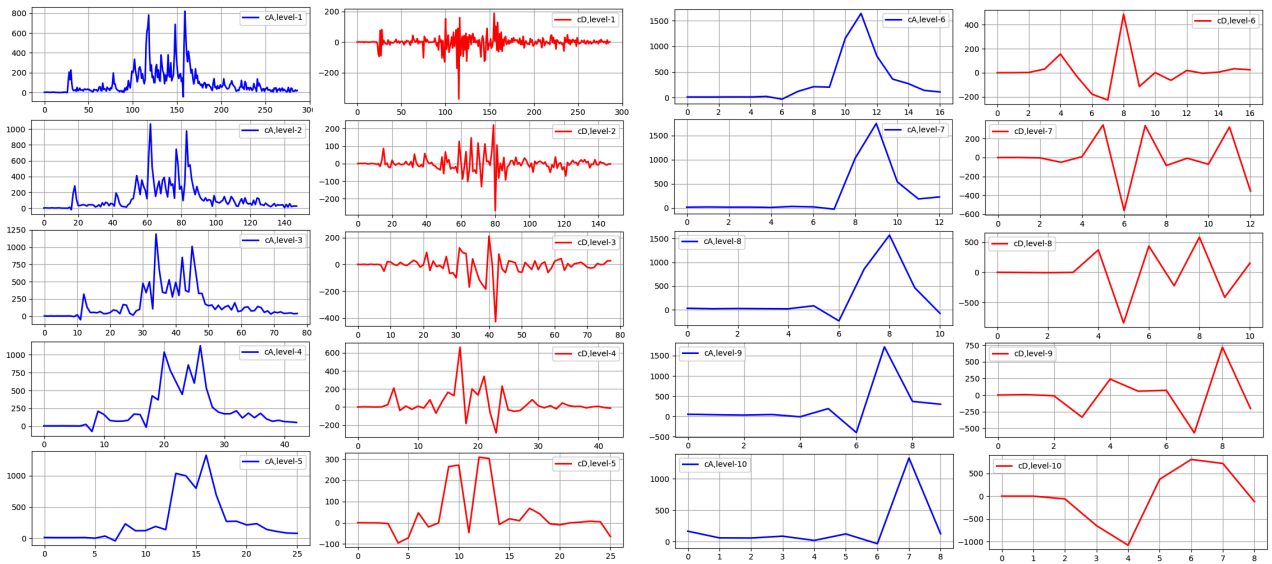


Fig. 22. Fifth-order Dobesi wavelet decomposition (1-10 levels). Russian news

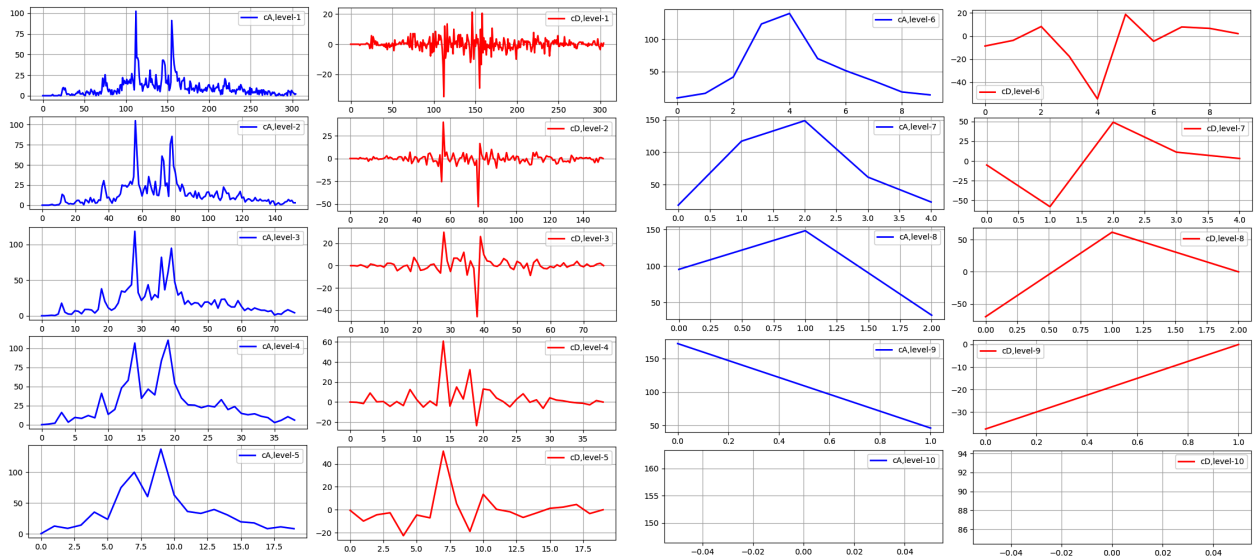


Fig. 23. First-order Dobesi wavelet decomposition (1-10 levels). SRDLR news

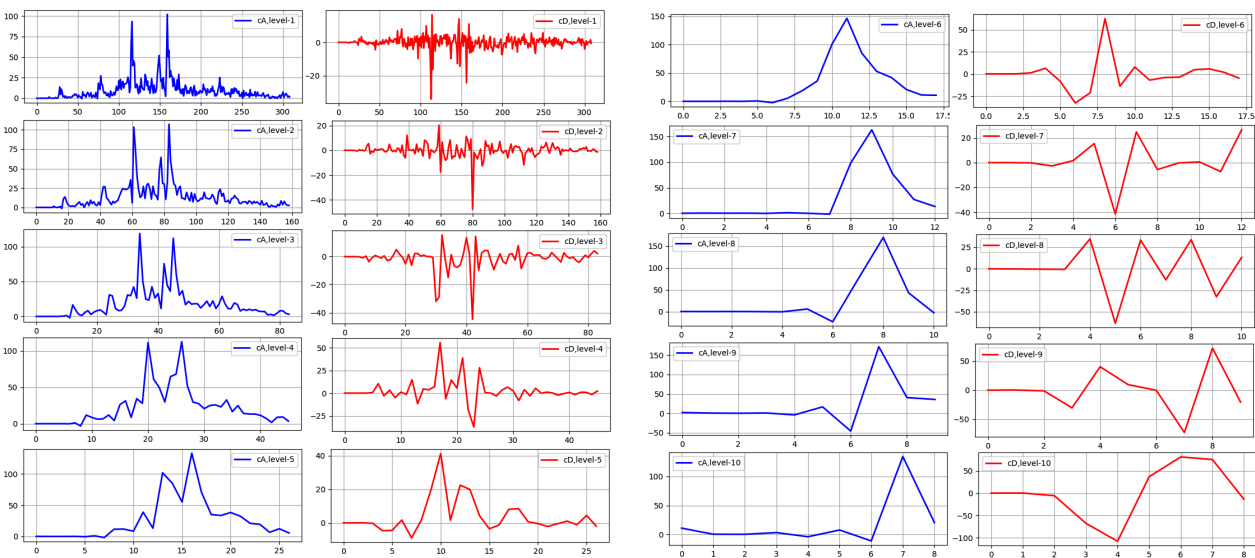


Fig. 24. Fifth-order Dobesi wavelet decomposition (1-10 levels). SRDLR news

relates with approximately equal emotional news shares in this region. Can be assumed that for this region news bursts are characterized by both positive news and negative.

Dynamics of all news almost completely coincides with the dynamics of negative news in SRDLR region. This may be partly due to the fact that negative news value is more than 57 percent. At the same time, on each decomposition level there are certain patterns inherent in a number of news in Russian region. Therefore, the final conclusion is that much of the news in the region, both negative and positive, is characterized by outliers and can be considered as information influences.

Step 4: Factor structure analysis

Factor structure is a set of factors affecting the system formation and being system or process component. Conceptually, social systems structure is very similar to neural networks structure and both of them consist of interconnected elements set divided into certain hierarchical levels and directly and feedback interacting each other. Another similarity is the nonlinear relationship between the system elements, as well as the simultaneous parallel processing of information. Thus, latent structure, e.g. the number of neurons on the latent layer, can be interpreted as the number of factors acting in the social process or system.

Item factor analysis (IFA) is a popular method for summarizing a number of categorical item responses using a smaller number of continuous latent variables. It is an indispensable tool for item analysis as well as test construction and scoring in psychological and educational measurement research [5, 8]. To discover new patterns from a huge database by knowing what factors affect the system and what information should be extracted is an essential step in data mining. The more precisely a factor analysis is done, the better the performance of the clustering and the classification in pattern recognition becomes. In factor analysis, the selected factors should be significant and complete. The necessary condition of significance requires the selected factors to be independent and important while the sufficient condition of completeness would provide complete information [5].

Experiment details

During the experiment direct distribution architecture (FNN) network and its *MLPRegressor* implementation on Python programming language from the *sklearn* library were used.

Feedforward Neural Network is a class of no feedback networks. Signals in such networks propagate only in one from the input neurons to the output through the hidden layer direction. FNNs are mainly multilevel classifiers of logistic regression and also called Multi Level Perceptrons (MLP). In this mode hidden layer is required. Thus, the use of several such layers makes sense only when using nonlinear activation functions.

The following 3 time series experiments were conducted:

- 1st: Ukrainian news number prediction, knowing news amount in Russia and SRDLR. Data from Russia and SRDLR were taken as X data. Ukraine data are Y. Network architecture is presented with 1 inner layer consisting 6 neurons.
- Russian news number prediction, knowing news amount in Ukraine and SRDLR. Data from Ukraine and SRDLR were taken as X data. Russian data are Y. Network architecture is presented with 1 inner layer consisting 6 neurons.
- SRDLR news number prediction, knowing news amount in Ukraine and Russia. Data from Ukraine and Russia were taken as X data. SRDLR data are Y. Network architecture is presented with 1 inner layer consisting 6 neurons.

Performed experiments results are given in table 4:

Table 4. Factor structure analysis results

	1 Ukraine ru	2 Russia ua	3 SRDLR ru
Neural network accuracy	0.85	0.88	0.75
Hidden layer neurons number	6	6	6
Hidden layer neurons weights	0.46 1.15 1.16 0.58 0.22 -0.44	0.71 0.43 -0.54 -0.71 0.38 0.83	-0.56 0.53 0.09 0.10 -0.22 0.86
Edges weights from X_1 to hidden layer neurons	1.06 1.03 1.14 0.96 0.77 -0.58	-0.37 -0.08 -0.05 -0.12 -0.32 0.2	-0.06 0.043 -0.3 -0.18 -0.48 -0.05
Edges weights from X_2 to hidden layer neurons	-0.24 0.99 0.39 -0.12 0.07 0.45	0.54 0.34 0.06 0.97 0.76 0.92	0.59 -0.12 -0.2 -0.2 0.71 0.50

Due to experiment there are 6 factors in each system. The conclusions are following:

- 1) Forecasting accuracy for Ukraine and Russia is 85 and 88 percent respectively which means easier possibility to predict news in Ukrainian and Russian regions.
- 2) SRDLR region news are more complicated for prediction as direct Ukrainian and Russian information influences reflection and can be considered as fake spreading region under both Ukrainian and Russian impact.
- 3) Russian and SRDLR news have a very similar structure, and their values of the correlation entropy are lower than the values on the Ukrainian news.
- 4) Experiment proves in general the hypothesis about two information influences spreading models with same character and properties, but different behavior. First one is aggressive and aimed at influences spreading in order to achieve such interests, the second is passive and responds to attacks in order not to suffer damage.

- 5) It can be concluded that 3 news bursts in the Russian region were symmetrically reflected in the Ukrainian news sources. Hypothetically, it can be assumed that bursts are caused by information influences or fake news spreading.

Conclusions

An iterative approach to the information influence model using EDA, fractal and wavelet analysis were used to conduct a aimed results:

- 1) Data statistic characteristics identification and data outliers description has been provided by Exploratory data analysis. Collected results are interpreted with box plots.
- 2) Fractal analysis performing provided Hurst exponent and corellation entropy values. Social processes have been studied for both persistence and uncertainty.
- 3) Wavelet analysis was performed in order to perform patterns highlight and news process decomposition at different levels.
- 4) In the constructed model conditions the optimum factor number has been found due to influencing factor data structure allocation.

The paper presents a method based on wavelet and fractal analysis which can be used for implementation within the specialized media monitoring systems of unstructured text flows intellectual analysis. Theme and development in this area are extremely relevant for Ukraine today due to actively developing sources that retransmit the Russia's agenda, and work to destabilize the situation in Ukraine, messages from whose messages are disseminated by provokers publishing posts and comments, which are created for spreading Russian indoctrination in social networks, as well as comments under news from online news outlets, etc.

Forehanded use of such methods in sphere of security and national security would help to automate the

detection process of fake news distribution in initial phases, allowing experts of security and defense structures to take timely measures and prevent manipulative messages spread.

References

- [1] K. Shuy, A. Slivaz, S. Wangy, J. Tang, and H. Liuy, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, 9 2017.
- [2] S. Lorent, "Fake news detection using machine learning," 2019.
- [3] D. Lande, I. Subach, and Y. Boyarinova, *Intellectual data analysis theory and practice fundamentals in the cyber security field*. Institute of Special Communication and Information Protection National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, 2018. Manual.
- [4] V. Gorbulin, O. Dodonov, and D. Lande, *Information operations and society security: threats, counteraction, modeling*. Kiev: Intertechnology Publishing House, 2009. Monograph.
- [5] H.-F. Wang and C.-Y. Kuoy, "Factor analysis in data mining," *An International Journal Computers and mathematics with applications*, no. 48, 2004.
- [6] W. Ebeling, L. Molgedey, J. Kurths, and U. Schwarz, "Entropy, complexity, predictability, and data analysis of time series and letter sequences," *The Science of Disasters*. Springer, Berlin, Heidelberg, 2002.
- [7] I. Dobesi, *Ten lectures on wavelets*. Izhevskv: SRC «Regular and chaotic dynamics», 2001.
- [8] C. J. Urban and D. J. Bauer, "A deep learning algorithm for high-dimensional exploratory item factor analysis," 1 2020.