MENDEL
Soft Computing Journal

UNCONVENTIONAL METHODS IN VOYNICH MANUSCRIPT ANALYSIS

Ivan Zelinka[1,✉], Oldrich Zmeskal[2], Leah Windsor[3], Zhiqiang Cai[3]

[1]Department of Computer Science, Faculty of Electrical Engineering and Computer Science VŠB-TU, Czech Republic
[2]Institute of Physical and Applied Chemistry, Faculty of Chemistry, Brno University of Technology, Czech Republic
[3]Institute for Intelligent Systems, University of Memphis, USA

ivan.zelinka@vsb.cz[✉], zmeskal@fch.vut.cz, Leah.Windsor@memphis.edu, zhiqiang.cai.2016@gmail.com

Abstract

This paper discusses the possible use of unconventional algorithms on analysis and categorization of the unknown text, including documents written in unknown languages. Scholars have identified about ten famous manuscripts, mostly encrypted or written in the unknown language. The most famous is the Voynich manuscript, an illustrated codex hand-written in an unknown language or writing system. Using carbon-dating methods, the researchers determined its age as the early 15th century (between 1404-1438). Many professional and amateur cryptographers have studied the Voynich manuscript, and none has deciphered its meaning as yet, including American and British code-breakers and cryptologists. While there exist many hypotheses about the meaning and structure of the document, they have yet to be confirmed empirically. In this paper, we discuss two different kinds of unconventional approaches for how to handle manuscripts with unidentified writing systems and determine whether its properties are characterized by a natural language, or is only historical fake text.

Keywords: *analysis, complex network, fractal geometry, text, Voynich manuscript.*

1 Introduction

This paper discusses the possible use of various unconventional algorithms and techniques on analysis and categorization of the unknown text. We selected the Voynich Manuscript (hereafter, VM) as the experimental text and quantity of interest, as it is the most famous illustration of a hand-written codex in an unknown writing system (amongst the such books as the *Book of Soyga, Codex Seraphinianus, Hypnerotomachia Poliphili, the Oera Linda Book, the Ripley Scrolls, the Smithfield Decretals, the Rohonc Codex, the Red Book, Prodigiorum Ac Ostentorum Chronicon*).

The timeline of the VM and its researching can be found in various research reports and publications as for example in [4], [7], [11], [13] or well written chronology in [33]. The VM is full of symbols that were partially interpreted by [30], and it is claimed that various plants, animals, a mineral, and cities were deciphered. These claims, however, have not been verified. Other scholars have also examined the codex's properties, including [7], [12], [25], [18], [21] [24], [26] or [15]. Unaffiliated researchers and hobbyists have also pursued investigations into the Voynich Manuscript. The VM studies are also available on the web as the web documents and the web blogs (not a scientific document) as for example [18], [28] or [31].

In this paper, we do not attempt to translate or decipher the codex. The aim of this paper is to identify statistically and computationally meaningful attributes of the text when analyzed by means of two unique methods. The structure of the paper is as follows. First, we provide a brief overview of VM. Next we describe our experimental design: an algorithm for fractal geometry ([10], by SW HarFA [20]). Using this methodology, we analyze the VM alongside the Ernst Hemingway book *The Old Man and the Sea*, and the Bible. We then describe the fractal properties of language and the complex networks of topology, including visualizations of the VM, the Bible, and *The Old Man and the Sea*, translated into 120 languages[1,2]. The aim of this workflow is to investigate how much VM differs from the books written in natural language. In the second experimental section, we perform an analysis based on the text conversion into the complex-like network ([19], [9], Fig. 10 and Fig. 11), where each vertex represents one word from the manuscript. Also randomly generated text is used for such analysis. Our analyses demonstrate that VM is likely **not** a historical fake, but rather that it is written in natural, yet unknown/undecrypted/undecoded, language.

---

[1]https://www.bible.com/languages
[2]https://ebible.org/download.php

2 Voynich Manuscript

The Voynich manuscript [11], Fig. 1, is a mysterious illustrated book, written in an unknown language. According to the radiocarbon dating (C_{14}) it was written sometime between 1404-1438. The book is named after Wilfrid Michael Voynich (Wojnicz, Vojnic), an American merchant of Polish-Lithuanian origin who acquired the manuscript in Italy in 1912. At present, the text is owned by Yale University under the designation MS408.

The manuscript is the subject of interest of many scholars of various disciplines, especially cryptologists. It is also connected with the Czech environment, as the first known owners were scholars from Prague, and it is even possible that it was written there. The book probably had about 272 pages in 17 sheets, but only 240 were preserved. The text is written on parchment with unknown lettering (20-30 characters with simple punctuation) in an unknown language. The manuscript consists of over 170,000 characters with spaces dividing the text into about 35,000 groups of varying length (accepted as words, or discrete lexical items) while 8114 of those words are considered unique "word types". In the entire book the characters and words distribution seems to be affected rules and indicates the existence of spelling and grammar. Statistical analysis, as in [1], [14], of the text suggests that Zipf's law applies to the language used, the average length of each word is roughly the same as Latin and English [32]. Importantly, the language used in Voynich's manuscript is quite different from the European languages. However, some signs suggest similarities with Semitic [or Indo-European?] languages, although some scholars have suggested that the manuscript is of Latin American origin [29]. Several words in the manuscript are also written in Latin, in a language reminiscent of ancient languages around the Mediterranean Sea, although it is possible that the Latin passages were added later. According to the human-generated classifications of the illustrations, the book is divided into several parts; it is likely that it is a tractate from several areas:

- Botanical part. The illustrations contain pictures of common European plants. However, most of them are difficult to identify; some have not yet been identified.

- Astronomical part. Illustrations include astronomical diagrams, astrological diagrams, and symbols, sketches of the zodiac. In some diagrams, the names of the months of the year are Latin, in Romanesque, perhaps in Catalan or Occitan.

- Biological part. The illustrations show mostly miniatures of naked women, bathers in strange formations, some of which resemble body organs. Some women have crowns.

- Cosmological part. Illustrations remind remotely of maps of a strange landscape or cosmological sketch. There are castles, perhaps volcanoes.

- Pharmacological part. Illustrations show parts of plants (roots, leaves). Perhaps it is a pharmaceutical prescription, which may also suggest that the text is divided into short paragraphs in this section.

- Recipes. Without illustrations, the text is divided into short paragraphs, separated by blooms or stars.

The text has not yet been decrypted, and its history and provenance are very fuzzy. One way to determine the time of the book's creation is according to the style of hairstyles, clothing, and castles shown in the illustrations. A more scientific approach using radiocarbon dating methods was also used. The period of creation is estimated to be between 1404 and 1438. Most studies of the Voynich Manuscript have relied on qualitative analysis, using the illustrations as benchmarks to demarcate the chapters and contents. Given recent advancements in computational linguistics methodology, we are now able to explore this document using rigorous, statistical approaches.

3 Fractal Geometry and the Text

We now turn to the empirical analysis of the VM, using fractals. Fractals are geometric objects whose basic structure, fragmented or irregular, is repeated at multiple scales. The term, coming from Latin, was proposed by the mathematician Benoit B. Mandelbrot in the 70s in 20th. century, to describe the quality of these fractured geometrical objects. Fractal geometry describes self-similar objects, a characteristic feature of complexity and complex dynamics. Self-similarity is a phenomenon that appears in the dynamics of many natural, artificial or mathematical systems [16]. Self-Similarity is the property of a geometric object that contains a part which is identical, or very similar, to the geometric structure of the whole object. In other words, a subset of the parent object is similar to the parent object.

Mandelbrot's book, The Fractal Geometry of Nature [16], became a bestseller and described how to construct very complicated structures using simple principles. However, the history of this discipline also has deeper roots, starting in the 17th century with recursive self-similarity stated by the philosopher and mathematician Gottfried

Figure 1: A few page samples from Voynich manuscript.

Leibniz. Then, in the 19th century, the mathematical study of continuous, but not differentiable functions, became increasingly rigorous. It was Karl Weierstrass who first described a function that was continuous but without derivative at any of its points. Relevant researchers who also made contributions within this field were Georg Cantor, Gaston Julia, Henri Poincare and others [8]. Nowadays, fractals have several applications like data compression, antenna systems, modeling natural patterns, computer graphics, digital art, etc [3], [16].

Self-similarity, as a typical attribute of fractals, can be shown graphically on two classic fractal objects: a snowflake and fern in Fig. 2. Upon zooming in on any part of the object, its structure will resemble the basic object again.

There are many algorithms for constructing fractals, as for example:

- Iterated Function System (IFS) uses geometric replacement rules, that may be deterministic or stochastic [6].

- Escape-time fractals use a recurrence relation at each point in a certain space [3].

- Strange attractors use the solutions of a system of differential equations that exhibit chaos [8] and have a fractal dimension [8].

- Lindenmayer systems (L-systems) use string rewriting [23] and a formal grammar to resemble branching patterns present in many biological structures like plants, cells from the immune system, etc [23], [3].

The most known algorithms are the Iterated Function System (IFS) and the Escape-time algorithms that present a tight relationship [6] between them. An essential characteristic of affine transformations is that we need a few numbers for a perfect description of an object. Each fractal has its unique description via a few numbers (coefficients of affine transformations [6]), and a slight change over them also changes the shape of the fractal. On the other hand, Escape-time algorithms work differently as it is based on the idea of convergent and divergent trajectories over a search space. They take a point from such search space and calculate how many iterations are needed to leave a certain predefined area. The number of iterations is then used to determine the colour for such point [16]. Thus, IFS produce black-white pictures while TEA colored one.

Because fractal presence and attributes have been observed and discovered in any natural language as reported in [10], we employ this methodological approach for analyzing the VM in a similar way as reported in pioneer work [9] by L. Hrebicek, then in [17], [27], [5], [2].

4   Complex networks

The complex network is in fact a *graph*, i.e. a structure, that consist of nodes (e.g. vertices) and edges (connections amongst them). The topology is usually specified regarding regular lattices or complex networks [19], which in recent years has become one of the best-developed areas of complex systems theory. The structure of complex networks can be observed in many systems. Large-scale networks, exhibiting complex patterns of interaction among vertices exist in both natures and man-made systems like genetics, ecological ecosystems, economics, sociology, scientific publications, the Internet, power grids, among many others [19].

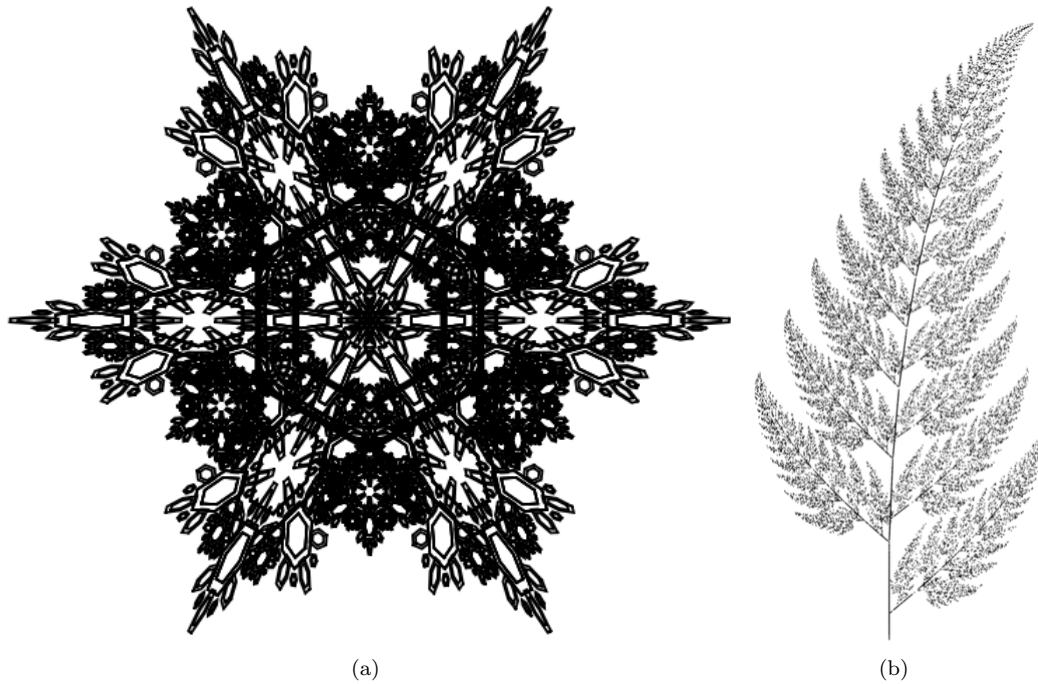(a)   (b)

Figure 2: Fractal samples, the snowflake, and the fern.

The name complex networks [19] comes from the fact that they exhibit substantial and non-trivial topological features, with patterns of connection between vertices that are neither purely regular nor purely random. Among many studies, two well-known classes of complex networks are the scale-free networks and the small-world networks (see examples in Fig. 10). Research in the field of complex networks has become interdisciplinary and had joined researchers from many scientific fields and perspectives. Complex networks are very relevant in the scope of this article, so we recommend to read its introduction in [19].

Complex networks' connection with analyzing text and literature was also introduced in [9] by L. Hrebicek, the Czech scientist and expert on the Turkish language. This work and idea have inspired us to use the complex network approach on the Voynich manuscript. The complex network visualization also inspired by ideas in [9] was used to convert the Voynich manuscript into a visual form for further analysis, see Fig. 10 and Fig. 11.

5 Experiment design

For the VM analysis was used fractal geometry and very basic complex network analysis. Experiment conditions and parameter setting is described here.

5.1 Data

The data used for our experiments were based on Hemingway's novel *The Old Man and the Sea* (120 language mutations), the Bible book in English (at the moment we are expanding those experiments for Bible in 150 languages) and computer translation of the Voynich manuscript [3]. As the Voynich text was used the computer translation known as the *Currier*, while other existing as *Benninett, FSG, Frogguy, Eva, Takahashi_EVA*, ... has not been used here. These computer translation - character sets differ in symbols that interpret individual letters from VM. Used data must be preprocesed. Original source files contained symbols and letters that were not a part of the text, so firstly before experiments all those objects were remover till only words from own text remains, see Fig. 3. During the analysis we found, as the curiosity, the most frequently used words, see Fig. 4.

5.2 Used software and hardware

For reported calculations, algorithms and visualizations have been used standard PC with an Intel Core I7 processor and 8Gb RAM with SW *Mathematica*® v. 11.0. Additionally for the fractal analysis was used SW HarFA [20] to calculate the fractal dimension of the text.

---

[3]http://www.voynich.nu

{FGAG2, GDAE, AR, GHAM, SOE, SORG, 0D0RC2, GDOR, SOE8G, 2ORG, DZAR, ORG, DAIR, THAM, SOR, AR, HZAR, HZAR, 8ALA, 2GAIIR, SCDG, OR, GDAM, SO8, HZOARG, HZC2, 8ARAM, 2G, 00M, OHCCG, OHCAR, ROEOHG, HZAAR, 8AM, ODAM, OR, ODAL, 2AIRG, TCAR, HZAM, PZAR, FZAM, G8ARAISG, O8AR, SG, SOE, ··· 33 571 ···, 4ODA2, AECCGTG, O4ODAM, AE, SCG, 4ODAR, ODARAE, ODCG, SPZCG, OHCCG, OODAR, ODG8G, O2AM, SDG, 4ORAM, TDZCG, 4ODCG, EDCTG, ODCCG, ODAE, TC8, DAEG, 2GDAR, AM, OEDACG, 8AM, OEG, 0A2OR, SCG, 4ODCG, ODCOEAL, TCG, 4OE, OR, TCCG, 4OR, ARAK, OE, EDAL, 2O8AE, TAE, THZG, TDZG, 4OE, AM, ARG}

large output | **show less** | **show more** | **show all** | **set size limit...**

(a)

large output | **show less** | **show more** | **show all** | **set size limit...**

(b)

Figure 3: The Voynich data preprocessing, a) original data, b) the same data under font, a screenshots from the *Mathematica*® software.
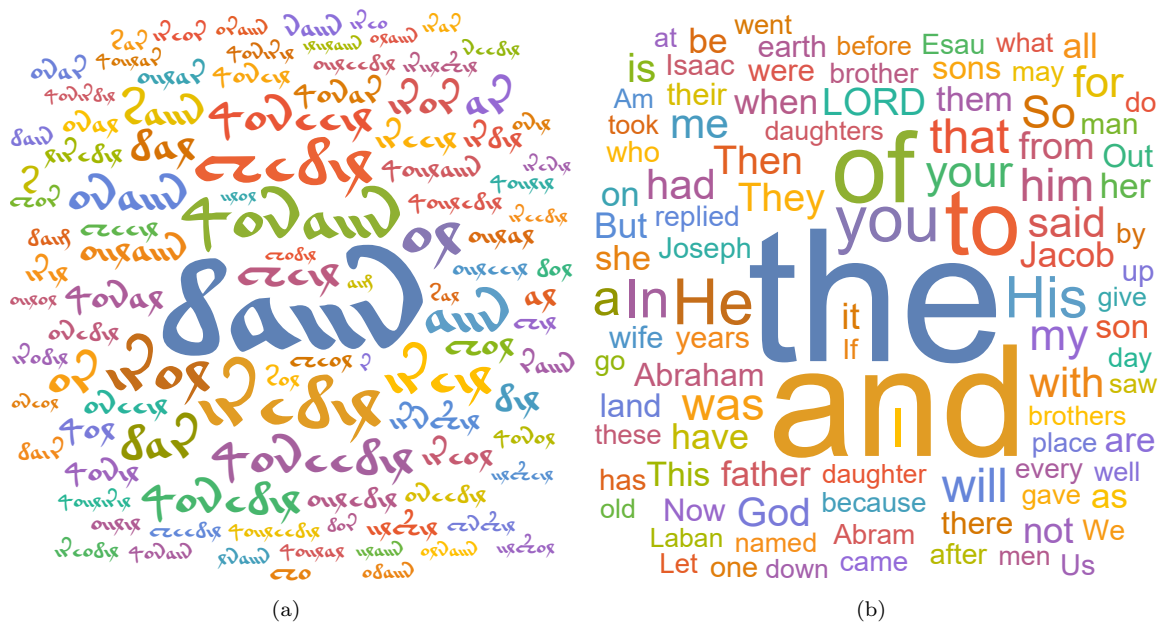
(a)

(b)

Figure 4: The most frequent words in Voynich manuscript (*8am* in Voynich computer code) and the Bible (*The*). The function *WordCloud* has been used in *Mathematica*®.
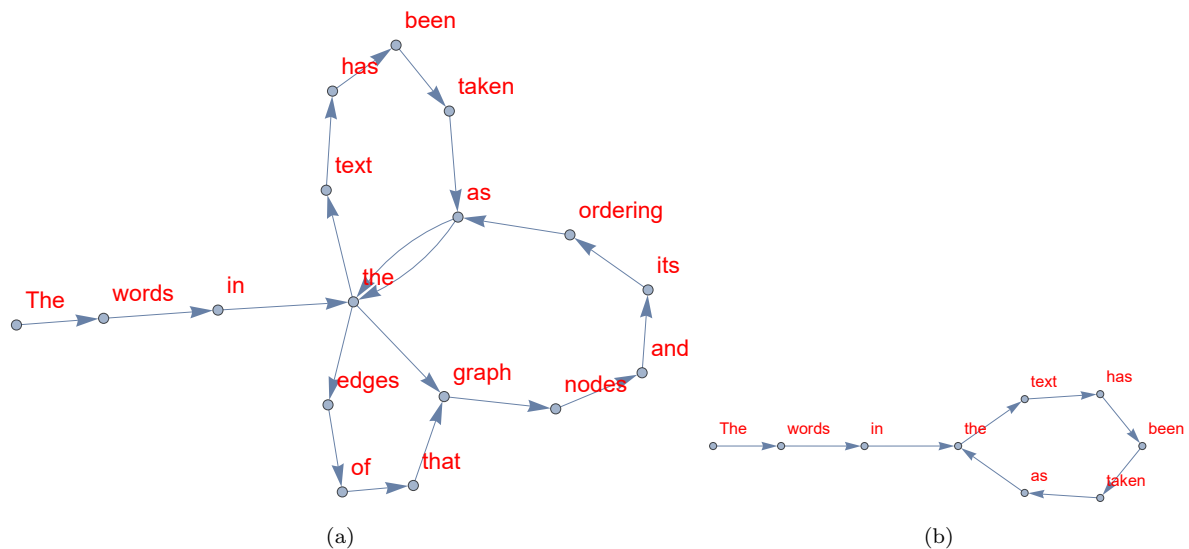
(a) (b)

Figure 5: The principle of the text-graph conversion.

5.3 The fractal analysis

A central question about the Voynich manuscript is its authenticity - whether or not it is a fake. The veracity of the Voynich Manuscript has been hotly debated, i.e., whether it is a randomly generated nonsense strings, or whether it is written in some natural language. For our analysis, it is not relevant if the VM is open or encrypted. To find the scientifically supported answer and investigate this question, we use fractal analysis and complex network visualization on the text of the VM.

For analysis of the Voynich manuscript, we used $Mathematica^{®}$ software and HarFA [20]. The $Mathematica^{®}$ software was used for the visualization of our data and also for all computation and analysis joined with compex network approach. The HarFA has been used for fractal analysis. In the fractal analysis, we calculated the fractal dimension of the Voynich manuscript and the Hemingway's novel *The Old Man and the Sea*, across 120 languages. In the second experiment, we also use the Bible in English translations.

5.4 The complex network visualization

In the case of a complex network, visualization was used only Voynich manuscript itself and compared with randomly generated text and its visualization as the network, to show that the fake text, based on random word writing, has different network structure and network parameters. Thus among the many of our various experiments, fractal analysis and complex network analysis and visualization are reported here. The conversion of the VM into a complex network has been done in a simple manner. The words in the text have been taken as the graph nodes and its ordering as the edges of that graph. Simply, if the text is *The words in the text have been taken as the graph nodes and its ordering as the edges of that graph*, then it is taken as the $The{\to}words{\to}in{\to}the{\to}text{\to}have{\to}been{\to}taken{\to}as{\to}the{\to}graph{\to}nodes{\to}and{\to}its{\to}ordering{\to}as{\to}the{\to}edges$ ${\to}of{\to}that{\to}graph$. The symbol $\to$ is taken as the oriented edge from $Node_1$ to $Node_2$. The graph then looks as on Fig. 5. Such kind of visualization can capture important text features in a complex network structure. For example at Fig. 6 is visible that 110th word in the text is repeated, so there are two the same words. Because of the limited number of words in the VM, the line of the graph starts to create loops – nodes that become a local or global centre of the network. For example, the English Bible translation is one of the dominant nodes representing word *the*). An example of the VM network growth is in Fig. 7 and of The Bible at Fig. 8.

The word *the* is in a class of lexical items called function words [22]. Function words are the glue that binds content words together in sentences. These include articles, conjunctions, temporal and causal connectives, and prepositions. We may hypothesize that the most frequently occurring words in the VM - as in other natural languages - are function words.

The same has been done for the randomly selected words in order to simulate random-like-fake text. Two cases of randomly generated text were used. In Case 1, 500 words were randomly taken and converted to a network. The result is in Fig. 12. In the Case 2, to keep the experimental conditions as precise as possible, we generated a set of 7391 unique words (Set_1), and then randomly created a set of 33663 words (Set_2) from Set_1. Then the same text-complex network was generated. This part was repeated $100\times$, and networks generated using this method are shown in Fig. 11. To compare VM with a well-translated book, we also selected *The Bible*. The Bible network is shown in Fig. 8.
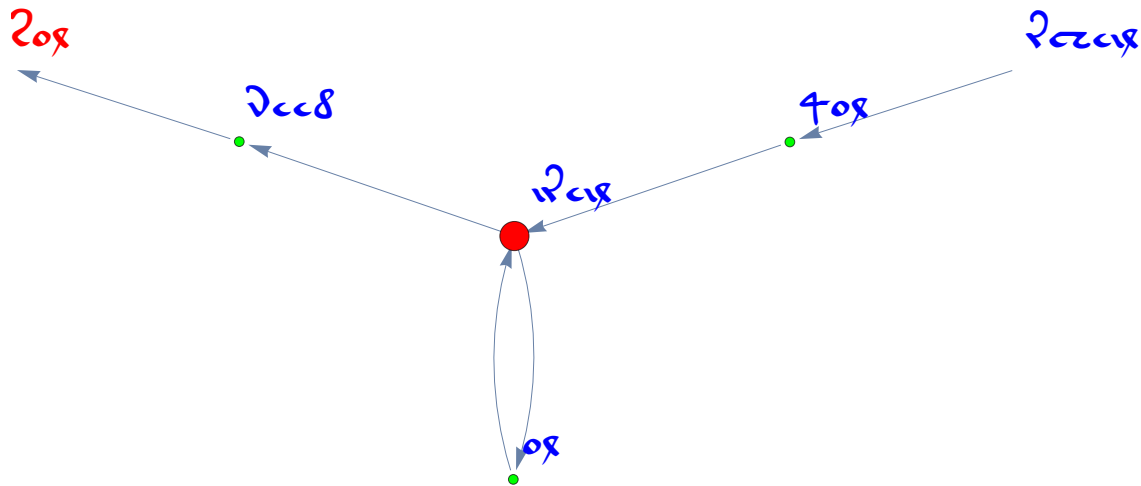
Figure 6: Word repetition. The color represent different degree centrality.

In order to compare properly in some global view the complex networks attributes from random, VM, or Bible networks, we generated and used the following Algorithm 1.

Input:
TEXT: the text (Bible or Voynich);
SSW: the max size of the sliding window;
RW: the range of the sliding window;
$y = 10, x = 1$: the counters;
while $y < SSW$ **do**

  • TEXT_y = Take first y words from position x ;

  foreach $TEXT_y$ and $x < RW$ **do**

    • CN_y = convert TEXT_y into a complex network as on Fig. 10 ;

    • Calculate degree centrality of the CN_y;

    • BDC = Select the biggest degree centrality from CN_y;

    • Plot BCD as color on coordinates (x,y) as in Fig. 13;

    • x++;

  end

  • y++;

  • x = 1;

end

Algorithm 1: General building pseudocode for the text degree centrality map as on Fig. 13.

The Algorithm 1 simply construct networks from the text of different length (see Fig. 13 y-axis from 10 to 200 words); for each text window size (string) is that done for positions from x=1 to 1000. For each network we calculated degree centrality [4] and represented by color in the figure, which can be considered as the map of the degree centrality of the text over the specific text range. The number of word strings processed for those maps at Fig. 13 was 200,000 (200×1000) for each and 5,000,000 (1000×5000) for each Fig. 15. A **10,400,000** word sets from VM was processed to get degree centrality maps at Fig. 13 and 15.

---

[4]https://en.wikipedia.org/wiki/Centrality
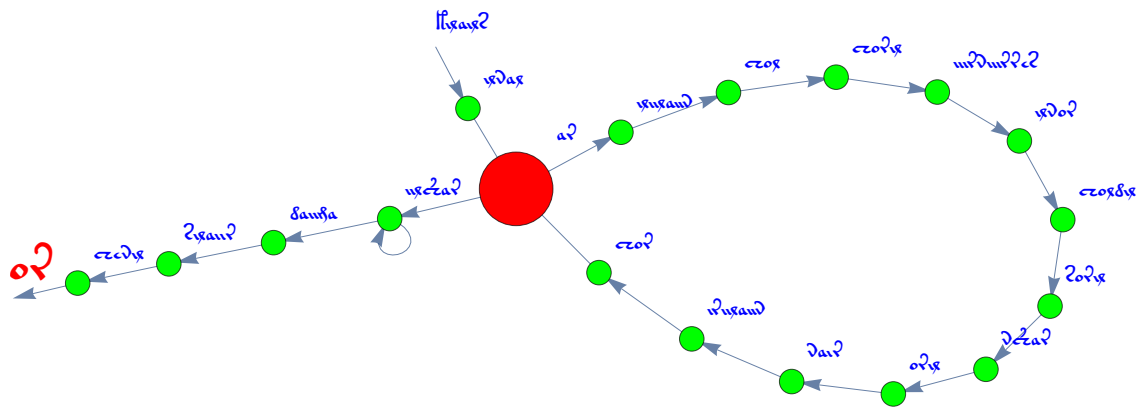
(a)
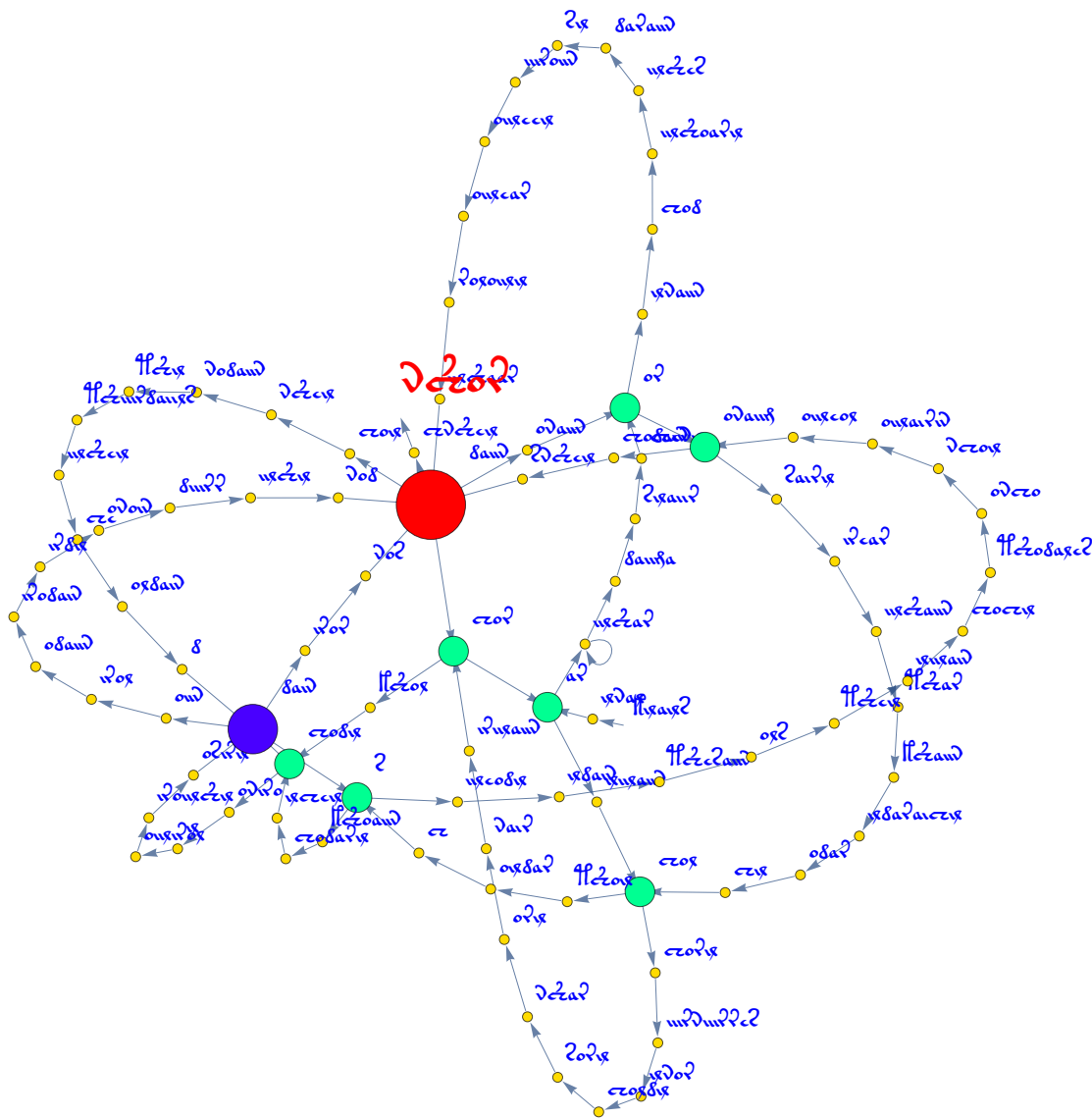
(b)

Figure 7: The growth of the Voynich network. The color represent different degree centrality.

Figure 8: The Bible network, first 500 words - an example.
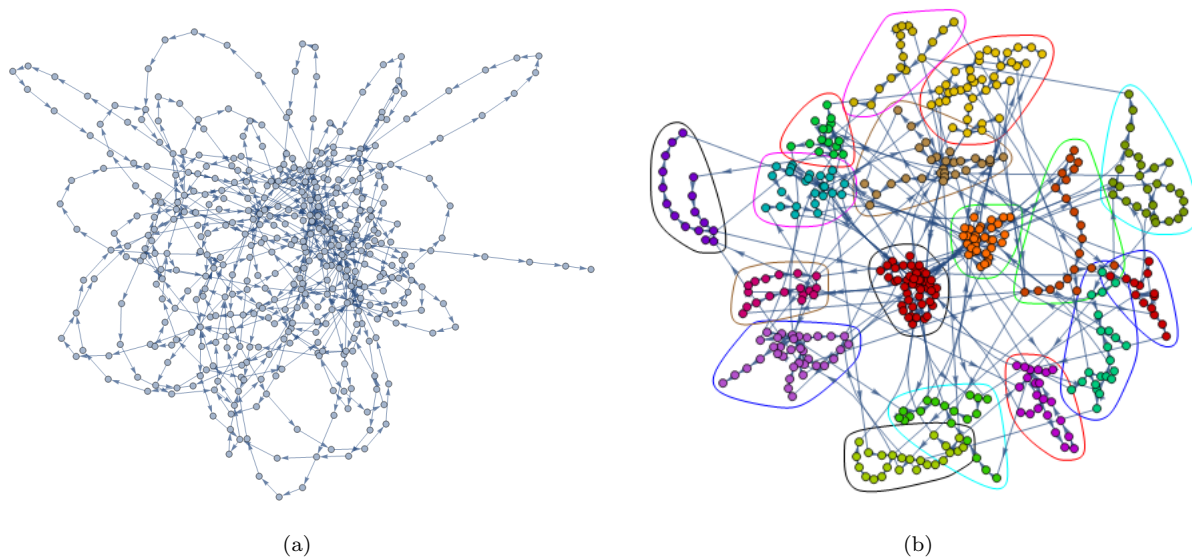

Figure 9: The Fractal analysis of the Voynich manuscript (red dots) vs. Hemmingway's *The Old Man and the Sea*, across the 120 language mutations. On the x and y axes are different fractal dimensions of the text.

(a)   (b)

Figure 10: The complex network-like visualization based on a 500 words from the Voynich manuscript (left) and the community (right)

6   Results and Discussion

Preliminary results are discussed here. Firstly we briefly show fractal analysis and then analysis on textual networks.

6.1   Experiment 1: Fractals

The first experiment focused on the fractal analysis of the Voynich manuscript. The HarFA [20] software was used to calculate fractal dimension [16], [10] of the text and compared with Hemingway's novel *The Old Man and the Sea* that in 120 language versions. The results are visualized in Fig. 9. The grey dots are related to Hemingway's novel The Old Man, and the Sea in various languages and the red dots belong to the fractal analysis of Voynich manuscript. Figure 9 shows the fractal dimension as calculated by HarFA. In all our results of the analysis of the Voynich Manuscript, very similar graphs were obtained. The important fact can be deduced from those figures: the Voynich manuscript belongs to the class of natural languages because it lies inside the grey cloud representing Hemingway's novel *The Old Man and the Sea* in various natural languages, not artificial ones.

The x and y-axes are coefficients of a linear regression of the fractal dimensions (see Fig. 9 for an example). The red dot is the Voynich manuscript and the grey Hemingway's *The Old Man and the Sea* in other languages. The dots near to red one can be strong candidates for an original, natural language that was likely used to write the Voynich manuscript.

6.2   Experiment 2: Complex Networks

The second experiment was focused on the conversion of the Voynich manuscript into a complex network structure. Any logical text that carries some information, consists of letters, words and sentences that cannot repeat in an arbitrary order. Rather, due to the grammatical rules of the language, some repeated structures and text causalities must be observable. For example, some languages follow a subject-verb-object grammatical pattern, while others use subject-object-verb. If the text is converted into a network structure, then such a network must look different than a random-based. Its network attributes should also be different. In Fig. 10 we see the first 500 words from the Voynich Manuscript, and its sequence converted into a network (left). On the right is its community visualization. The same for random text, see Fig. 11, is done. Compare it visually also with Fig. 12. Based on calculated degree centrality also its map over the selected VM text was created (see Algorithm 1), as described above and visualized at Fig. 13, 14, 15.

These maps, as for example Fig. 13 can be understood so that with the increasing length of the sliding window over the text, more complex networks are revealed. Also, for windows, with constant length, those networks show different complexity when sliding from position 1 to 1000 (or 5000). This is not true for randomly generated texts as visible when compared Fig. 14 and Fig. 13 (Fig. 15 respectively).
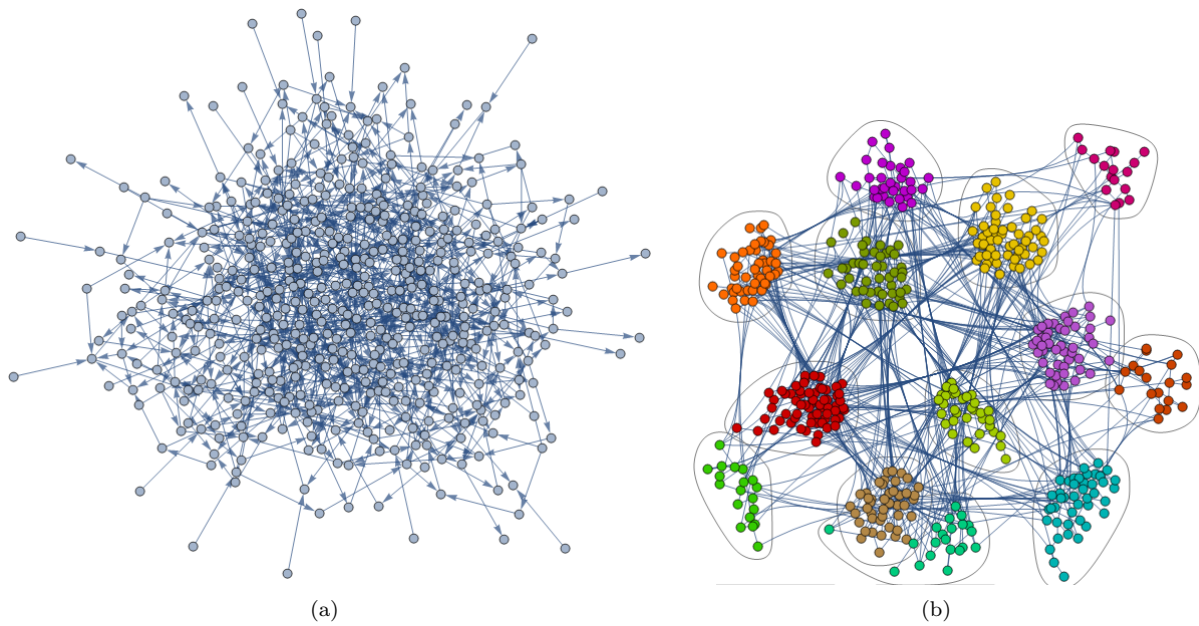
Figure 11: The complex network-like visualization based on a random text with 500 words (left) and the community (right), 500 nodes, Case 1.
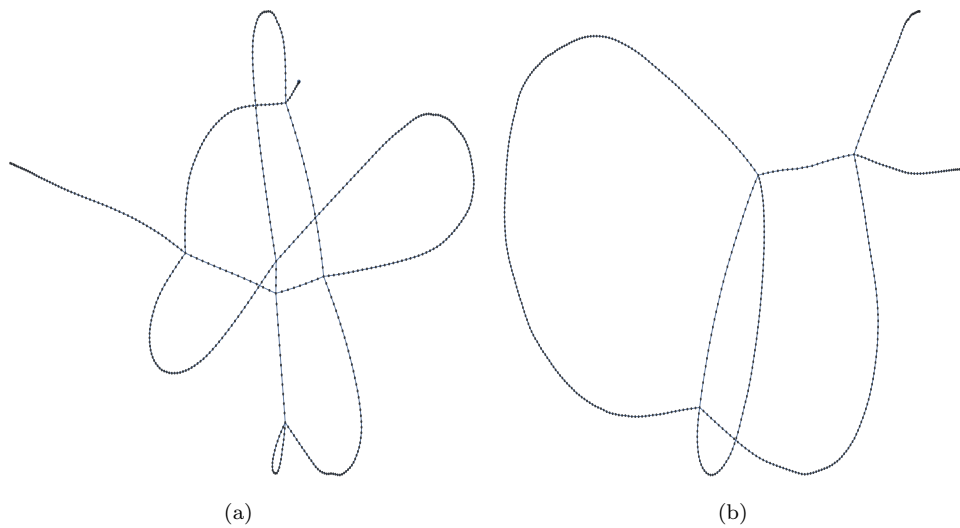

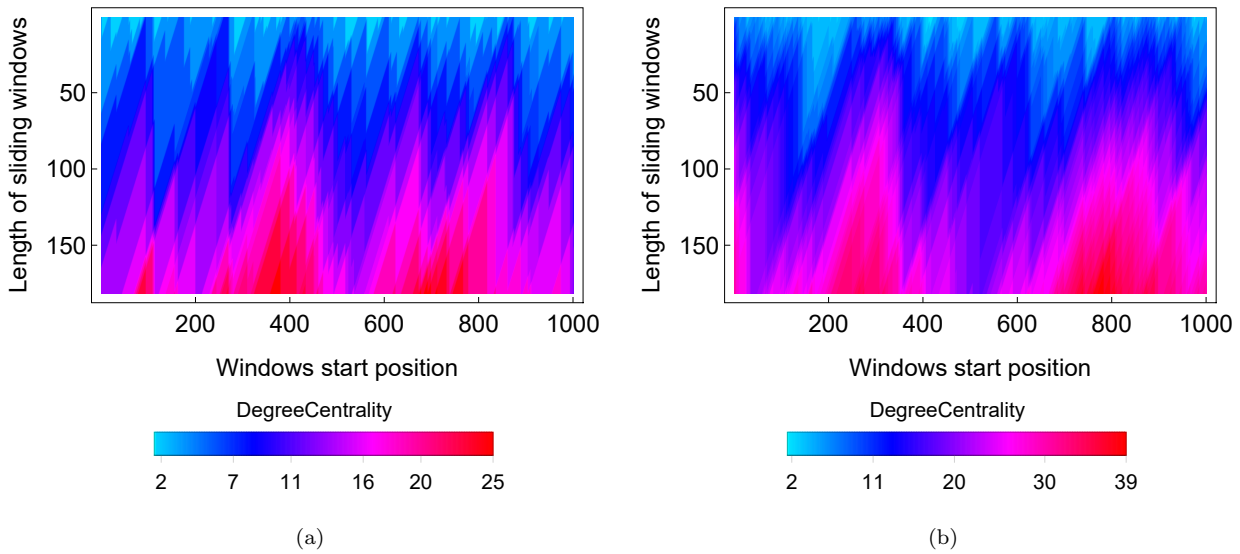Figure 12: Randomly generated networks, 500 nodes, Case 2.

Figure 13: Degree Centrality map of the **a)** Voynich manuscript. The set of words (used to created CN and calculate degree centrality) varies from 10-200 and shifted from the first word to the word on position at 1000. **b)** Bible in English. The set of words (used to created CN and calculate degree centrality) varies from 10-200 and shifted from the first word to the word on position at 1000.
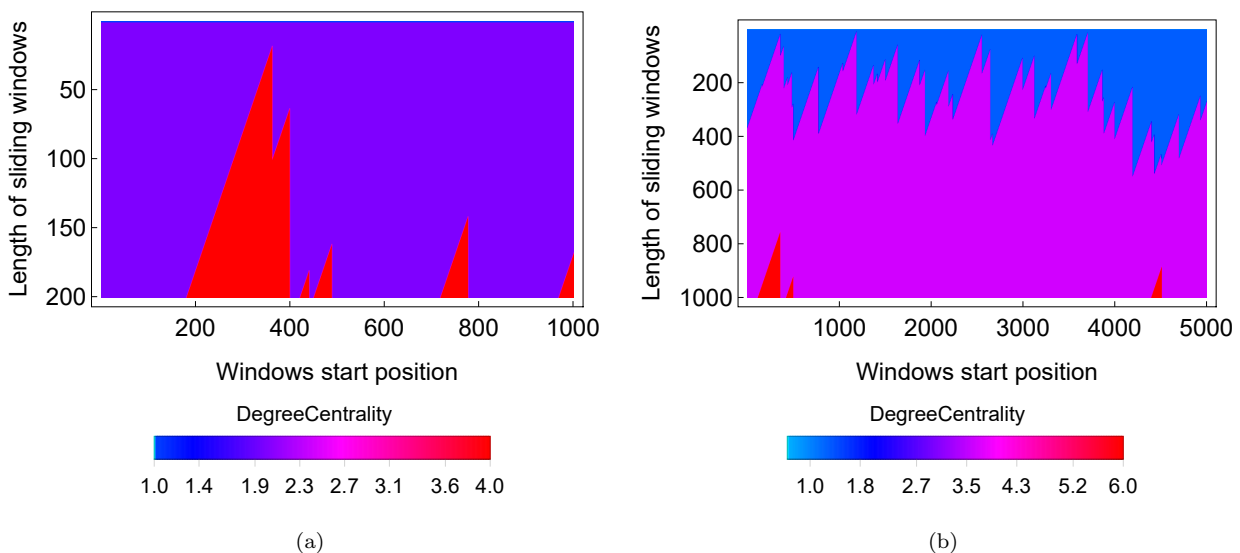

Figure 14: Degree Centrality map of the random text of Case$_2$. The set of words (used to created CN and calculate degree centrality) varies from a) 10-200 and shifted from the first word to the word on position at 1000, b) 10-200 and shifted from the first word to the word on position at 5000.
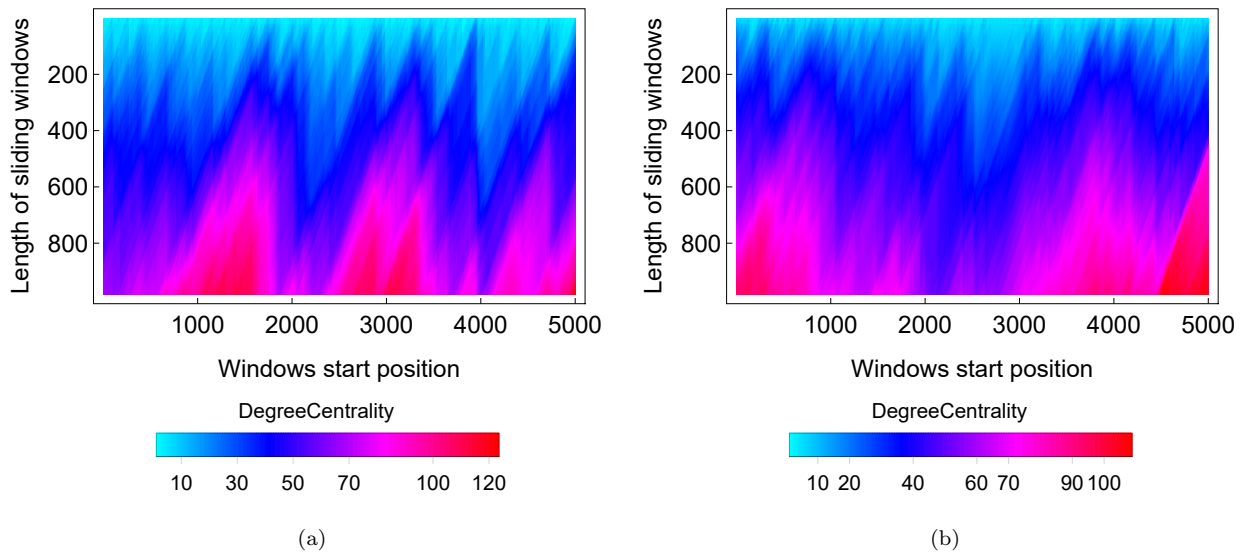
Figure 15: Degree Centrality map of the **a)** Voynich manuscript. The set of words (used to created CN and calculate degree centrality) varies from 10-1000 and shifted from the first word to the word on position at 5000. **b)** Bible in English. The set of words (used to created CN and calculate degree centrality) varies from 10-1000 and shifted from the first word to the word on position at 5000.

As it is from Fig. 10 and 11 visible, the random text creates more homogeneous graphs, which are on the first look different from graphs based on natural languages. This is also true when numerical analysis of attributes (i.e., various network centralities) of such networks is done. The fractal results, as well as complex network analysis, indicate the same. VM is very likely non-fake text, written in the natural language.

7 Conclusion

In this paper is reported our research on so-called Voynich manuscript with the aim to analyze whether this still undecoded manuscript is written in natural language, or whether it is just a historical text-fake from middle age authors. We have used HarFA [20] software for fractal analysis of the Voynich manuscript and as preliminary research (open and still in process), its conversion into a complex network structure [10, 11]. Our results indicate, that Voynich manuscript is likely written in some natural language because a) its fractal dimension set it among the Hemingway's novel *The Old Man and the Sea*, analyzed in the same way in 120 languages, b) complex network maps, generated from Voynich manuscript is significantly different from complex networks maps based on random texts. Thus our conclusion is, that Voynich manuscript is the text containing reasonable information rather than historical fake. As the future research, we are continuing in the fractal analysis, complex network analysis and another method used to aim better analyze this famous manuscript.

References

[1] Amancio, D. R., Altmann, E. G., Rybski, D., Oliveira Jr, O. N., and Costa, L. F. 2013. Probing the statistical properties of unknown texts: application to the Voynich manuscript. *PLoS One* 8, 7, No. e67310.

[2] Andres, J., Benešová, M., Kubáček, L., and Vrbková, J. 2012. Methodological note on the fractal analysis of texts. *Journal of Quantitative Linguistics* 19, 1, pp. 1–31.

[3] Barnsley, M. F. 2014. *Fractals everywhere*. Academic press, USA.

[4] Brumbaugh, R. S. 1974. Botany and the Voynich "Roger Bacon" Manuscript Once More. *Speculum* 49, 6, pp. 546–548.

[5] Chaabouni, A., Boubaker, H., Kherallah, M., Alimi, A. M. and El Abed, H. 2010. Fractal and multi-fractal for arabic offline writer identification. In *2010 20th International Conference on Pattern Recognition*. IEEE, pp. 3793–3796.

[6] Demko, S., Hodges, L., and Naylor, B. 1985. Construction of fractal objects with iterated function systems. In *ACM SIGGRAPH Computer Graphics*. ACM, Vol. 19/3, pp. 271–278.

[7] d'Imperio, M. E. 1978. *The Voynich manuscript: An elegant enigma.* Technical report, National Security Agency/Central Security Service.

[8] Hilborn, R. C., et al. 2000. *Chaos and nonlinear dynamics: an introduction for scientists and engineers.* Oxford University Press, UK.

[9] Hřebíček, L. 1971. *Turkish grammar as a graph.* Academia, Czech Republic.

[10] Hřebíček, L. 1998. Language fractals and measurement in texts. *Archív orientální* 66, 3, pp. 234–242.

[11] Kennedy, G. and Churchill, R. 2004. *The Voynich manuscrip.* Orion Publishing Company, London, UK.

[12] Kennedy, G. and Churchill, R. 2006. *The Voynich manuscript: The mysterious code that has defied interpretation for centuries.* Simon and Schuster, New York, USA.

[13] Kircher, F. and Becker, D. 2012. *Le manuscrit Voynich décodé.* Ed. le Temps présent, France.

[14] Landini, G. 2001. Evidence of linguistic structure in the Voynich manuscript using spectral analysis. *Cryptologia* 25, 4, pp. 275–295.

[15] Levitov, L. E. 1987. *Solution of the Voynich manuscript: A liturgical manual for the enduring rite of the Cathari heresy, the cult of Isis.* Aegean Park Press, Laguna Park, USA.

[16] Mandelbrot, B. B. 1983. *The fractal geometry of nature.* WH Freeman, New York, USA.

[17] Martin, J. R. 1995. Text and clause: Fractal resonance. *Text-Interdisciplinary Journal for the Study of Discourse* 15, 1, pp. 5–42.

[18] Montemurro, M. A. and Zanette, D. H. 2013. Keywords and co-occurrence patterns in the Voynich manuscript: An information-theoretic analysis. *PloS one* 8, 6, No. e66344.

[19] Newman, M., Barabasi, A.-L., and Watts, D. J. 2011. *The structure and dynamics of networks.* Princeton University Press, USA.

[20] Zmeškal, O. 2019. HarFa sowtware. `http://www.fch.vut.cz/lectures/imagesci/includes/harfa$_$download.inc.php` , [Accessed: 2019-05-7].

[21] Pelling, N. 2006. *The curse of the Voynich: the secret history of the world's most mysterious manuscript; the intriguing story of the people, places, and politics behind the enigmatic "Voynich Manuscript".* Compelling Press, UK.

[22] Pennebaker, J. W. 2011. The secret life of pronouns. *New Scientist* 211, 2828, pp. 42–45.

[23] Prusinkiewicz, P. and Lindenmayer, A. 2012. *The algorithmic beauty of plants.* Springer Science & Business Media, Berlin, Germany.

[24] Rogers, R. C. 2007. The resilience of Aztec women: A case study of modern Aztec myths. *Journal of Humanities and Social Sciences* 1, 2. DOI: 10.1.1.555.4173

[25] Skinner, S., Prinke, R. T., and Zandbergen, R. 2017. *The Voynich Manuscript: The World's Most Mysterious and Esoteric Codex.* ReadHowYouWant, US.

[26] Smith, M. E. 1996. *The Aztecs.* Blackwell, Oxford, UK.

[27] Tang, Y. Y., Ma, H., Mao, X., Liu, D., and Suen, C. Y. 1995. A new approach to document analysis based on modified fractal signature. In *Proceedings of 3rd International Conference on Document Analysis and Recognition.* IEEE, Vol. 2, pp. 567–570.

[28] Tiltman, J. H. and Bacon, R. 1968. *The Voynich Manuscript: "The Most Mysterious Manuscript in the World".* Baltimore Bibliophiles, USA.

[29] Tucker, A. O. and Janick, J. 2016. Identification of phytomorphs in the Voynich codex. *Horticultural Reviews* 44, pp. 1–64.

[30] Tucker, A. O. and Talbert, R. 2013. *A preliminary analysis of the botany, zoology, and mineralogy of the Voynich manuscript.* Herb News, Austin, TX, USA.

[31] Velinska, E. 2013. *The Voynich Manuscript: Plant id list.* `ellievelinska.blogspot.com/2013/07/the-voynich-manuscript-plant-id-list.html` , [Accessed 5 June 2017].

[32] Windsor L., Cai Z., and Zelinka I. 2019. Linguistic Properties of the Voynich Manuscript. In *Memphis DATA – A Data Science Conference.* Fedex Institute of Technology, The University of Memphis, Memphis, TN, USA.

[33] Zandbergen, R. 2017. *The Voynich Manuscript (2004–2017).* `www.voynich.nu/` , [Access 6 Jun. 2017].