Network marks of montage in audio recordings

D G Portnyagin¹, E I Kravchuk¹, A I Trufanov², A S Ivanov², O G Berestneva³, A A Tikhomirov⁴

¹ Irkutsk Investigative Office in Irkutsk Region, Investigative Committee of Russian Federation, 9a, Klar Netchkin st., Irkutsk, 664039, Russian Federation
² Institute of Information Technology and Data Science, Irkutsk National Research Technical University, 83, Lermontova st., Irkutsk, 664074, Russian Federation
³ Division for Information Technology, Tomsk Polytechnic University, 30, Lenin ave., Tomsk, 634050, Russian Federation
⁴ CIS Collaboration Center, Inha University, 100 Inharo, Nam-gu, Incheon, 22212, Republic of Korea
⁵ troufan@gmail.com

Abstract. Very often forensic domain processes required an approval of authenticity in audio recordings presented as admissible evidence. Standard techniques to search for editing in audio materials are rather long and wearisome. The paper proposes a network platform as an effective instrument for solving the above mentioned problem. A specialized software tool was developed to transform given audio data into set of nodes and links according to the algorithms of natural visibility graph and horizontal visibility. A comparative analysis of the derived network structures was performed with the use of popular Gephi software product. The results demonstrate the first advances of network paradigm for detection of audio montage, in addition the examples of trivial signals of those point on possible existence of a marker - the metric that responds to sound recordings tampering.

1. Introduction

One of the tasks that correspond to phonoscopic forensic examination is detection of montage in audio recordings presented as admissible evidence [1]. Standard techniques to search for editing in audio materials are rather wearisome and time-consuming processes [2]. Considering that any audio signal can be represented as a superposition of harmonic signals of various frequencies [3], it is somewhat reasonable to study the effect of editing on each frequency and combination of multiple and non-multiple frequencies while testing traditional or developing approaches for audio material analysis. By identifying common patterns for one frequency or a superposition of two or more frequencies, both multiples of each other and non-multiples, synthesizing these patterns, one can interpolate those to determine the presence of tampering in any studied audio signal.

Interestingly, in the last two decades, such a tool as complex networks has been developed and utilized to study a variety of complex systems [4-7]. Moreover, a general approach to time series analysis with rendering their network structure [8-10] occurred to be valid for application to diverse dynamic systems, including audio signal flows.

A core of this methodology is a technique of mapping time series (TS) onto complex networks by which the information encoded in the time series is transferred to the network topological features. It should be perceived that the analysis of audio signal implies not only representation of an original
signal in the form of a network model but also subsequent analysis of principal network metrics, heterogeneity of which will reflect presence of editing in the studied signal. Interestingly, the network models applied to sound data are still Russian exotics in the study of sound information in general [11-12]. At the same time, in world research practice they are used for various kinds of audio data, including speech, music and sounds of various origins [13-17].

2. Methods and tools
The task of audio signal analysis was initially separated into two sub-tasks and corresponding steps: first, converting initial data (monophonic or stereo audio signals) into a network structure and, second, analysing the concomitant network structures in order to identify metrics and parameters that indicate presence of editing.

Concerning step #1 one should imply that nowadays experts have considered a variety of popular and dependable approaches for mapping time series onto network structures just to assess those thoroughly:

2.1. Visibility graph (VG) set
The authors of [8] presented an intuitively clear and effective computational technique, that uses so-called visibility one. The technique transforms a time series into a network with two principal algorithms of natural visibility graph (NVG) and horizontal visibility graph (HVG) also supported by many other clones of those (e.g. one that convert a multivariate time series into multiplex networks [18]). It is of value that the constructed networks portray underlying nontrivial properties of a TS or images onto their specific topologies.

2.2. Recurrence networks RN
Recurrence implementation provides a clear transformation of a time series into a plot [19]. The plot is matched to a binary matrix which might be considered as the entity similar to adjacency matrix related to a network [20]. This very complex network reflects similarities of the phase space states at different time intervals for a system.

Notably the paper [21] observed that an appropriate recurrence rate minimizes the influence of noise on global clustering coefficient (C) but not on average path length (L)
The first paper on converting TS into networks [22] by Zhang and Small further was followed with:

a) TF topological features (super-family structure)
The study [23] characterized the dynamics of the time series data in line with selected structural features of the complex network based on correlation among specific types of dynamics with specific prevalence in the motifs: the adjacency matrix they considered differs from recurrence matrix and the properties they explored are topological specificities of the network.

b) Local sort (LS)
The work [24] gives a method to map TS into a weighted and directed complex networks. The proposed mapping technique differs significantly from the other existing techniques.

A set of segments is generated by a sliding window.

All generated segments are depicted with utilization of a doubly symbolic scheme represented by combination of absolute amplitude information and an ordinal pattern description.

In line with this construction, the given time series can be transformed into a network: segments with different symbol-pairs conform to network nodes and directed links shows the temporal succession between nodes. It is of significance that with this transformation, dynamics inherited TS is encoded into the network topology.

c) SGM (surrogate generation method)
The study [25] developed a surrogate generation method that represents natural and physical systems as complex networks. The method is close to that of a nonlinear TS analysis which uses a productive generation algorithm to get pertinent samples from the class of scale-free networks.
4) multiplex network-based sensor information fusion model (MSIF model)
This model [26] was proposed to study two-phase flow systems and their dynamics, so that the properties of heterophily and randomness where taken into account. MSIF model provides efficient fusion of multiplex signals to portray the complexity of the dynamics inherent to time dependent real systems.

2.3. CBS (coarse-graining based on statistics of segments method)

Coarse-graining based on statistics of segments (CBS) [27] was announced as the method to map time series into complex networks, the one that insensitive to diverse interferences. According to the method a slide window divides the flow into set of segments. Multi-scale entropy of the data is a factor to choose the width of the window.

The authors showed that CBS is still valid, even with strong disturbance of T, and demonstrate an essential robustness if compare to VG and Local Sort LS approaches.

Prior to study sensitivity of network models to interference and noise it is of reason to start in the forensic domain with simple and apprehensible instruments. Principally, the instrument should preserve encoded information containing in spatio-temporal records even being converted to network structures. And after mapping TS latent properties might be adequately clarified due to this preservation.

We chose a Visibility graph (VG) set as a key method to transform audio signals into networks.

Even Fortran 90/95 and Matlab codes for the visibility graph algorithms are available at http://www.maths.qmul.ac.uk/~lacasa/Software.html, one should imply the next considerations. Audio recordings that are under phonoscopic forensic examination to detect montage in those might last hours and days. Corresponding process of mapping such a long time series (with more than $10^6$ points) into complex network is time-consuming and requires high-performance tools which are sensitive to computational complexity of imbedded algorithms.

In case of HVG algorithm, at each step one safely assume confidently that no point after any other point of value larger than the current value will be horizontally visible from this current point. Such a consideration brings HVG time complexity equal to $O(n \log(n))$.

The work [28] noted that complexity of NVG algorithm if apply it directly, i.e. in line with [8] demonstrates a worst-case time complexity as $O(n^2)$, where $n$ is size of the series. This is explained easily as every pair of points in the time series should be checked on their visibility relation which leads to total number of checks equal to $n(n-1)/2$. Naturally such complexity becomes an issue for long time series with millions of points to be converted into large -scale networks. To overcome the problem, faster transformation algorithms based on Divide & Conquer strategy (DC) have been proposed to reduce the average-case time complexity to $O(n \log(n))$ for NVG [28].

Moreover a new method to compute visibility graph which employs Encoder/Decoder (ED) approach [10] is of significant interest per se. The ED scheme offers a computation that allows on-line assimilation of new data with no additional cost. However new data has not been so far an issue for phonoscopic forensic examinations.

In the current work a specialized software tool was developed to convert the studied audio data into tables of vertices and edges of the graph according to the algorithms of natural visibility graph and horizontal visibility graph. The tool comprises C# package that realizes the algorithm DC for NVG and straightforward approach for HVG (NVG-DC and HVG-S consequently).

3. Data
The second step of the network analysis envisages identification of metrics that indicate presence of editing. And just to prove detection of audio montage several signals containing one, two, four, and eight harmonics were selected as test examples. Thus the signals are:

- One harmonic of frequency 1000 Hz (Fig. 1).
- Two harmonics of frequencies 500 and 1000 Hz (Fig. 3);
- Four harmonics of frequencies 500, 1000, 1300 and 1900 Hz (Fig. 5);
- Eight harmonics of frequencies 500, 1000, 1300, 1500, 1900, 2000, 2300 and 2900 Hz (Fig. 7);
- A montage of signals of the above frequencies (Fig. 2, 4, 6, 8).

Figure 1. 1000 Hz signal.

Figure 2. 1000 Hz signal with montage.

Figure 3. 500 Hz + 1000 Hz signal.

Figure 4. 500 Hz + 1000 Hz signal with montage.

Figure 5. 500 Hz + 1000 Hz + 1300 Hz +1900 Hz signal.

Figure 6. 500 Hz + 1000 Hz + 1300 Hz +1900 Hz signal with montage.

Figure 7. 500 Hz + 1000 Hz + 1300 Hz +1500 Hz +1900 Hz + 2000 Hz + 2300 Hz + 2900 Hz signal.

Figure 8. 500 Hz + 1000 Hz + 1300 Hz +1500 Hz +1900 Hz + 2000 Hz + 2300 Hz + 2900 Hz signal with montage.
4. Findings
The presented test signals were transformed into graph (network) structures by the algorithms NVG-DC (Fig. 9) and HVG-S (Fig. 10). Then a comparative analysis of the derived network structures was performed with the use of popular Gephi software product [29]. In the analysis we collated the metrics, typical for an original signal and the metric for a concomitant modified signal. The network structures were visualized using "Yifan Hu" layout option.

Figure 9. Network structure derived by NVG-DC algorithm.

Figure 10. Network structure derived by HVG-S algorithm.

5. Discussion
A significant amount of computation made it possible to compare various statistical parameters and network metrics of tested audio signals. It should be noted that while converting the original signal into a network structure using the HVG-S algorithm, the "Eccentricity Distribution" metric provides the best visual representation of the montage (Fig. 11-18).

Figure 11. «Eccentricity Distribution» metric for 1000 Hz signal.

Figure 12. «Eccentricity Distribution» metric for 1000 Hz signal with montage.

It is important to emphasize that other metrics analysed in the study do not clearly identify the presence of tampering in the test audio signals.
Being applied the given approach to the analysis of signs of montage for more intricate audio signal such as musical fragment with embedded changes, the metrics "Eccentricity Distribution" do not provide unequivocal answer on presence of montage (Fig.19-20).
6. Conclusion

The results of this study demonstrated the first advances of network paradigm for detection of audio montage, also the examples of trivial signals of that point on possible existence of a marker - the metric that responds to sound recordings tampering. However, it is too early to claim that the problem has been solved, even within the limits of these simplest signals.

In reality for time series converted into a complex network by any of the above methods, noise, deformation, or falsification factors may significantly corrupt the topological structure of the inferred network. The studies [27] showed that as a matter of fact, the structure of complex networks gotten from raw record material definitely differs from that derived on the basis of corresponding distilled signals.

Network Science [4] provides the researcher with a powerful tool for analysis, but also requires a deep knowledge of the applied domain, creativity, ingenuity and non-standard methods of pre-processing data to create relevant ontologies, on the basis of which these data are transformed into network structures.

References

[1] Korycki R 2014 Authenticity examination of compressed audio recordings using detection of multiple compression and encoders’ identification. Forensic Science International 238 33-46 doi:10.1016/j.forsciint.2014.02.008

[2] Galyashina E I 2018 Diagnostics of the authenticity of digital copies of phonograms in phonoscopic examination: possibilities and limits of research Vestnik economicheskoi bezopasnosti 3 34-41

[3] Gonorovsky I S 1986 Radio engineering circuits and signals: Textbook for universities (Moscow: Radio and Communications)

[4] Vespignani A 2018 Twenty years of network science Nature 558 528–529 doi:10.1038/d41586-018-05444-y

[5] Boccaletti S , Latora V , Moreno Y , Chavezf M and Hwang D-U 2006 Complex networks: Structure and dynamics Physics Reports 424 175–308 doi:10.1016/j.physrep.2005.10.009

[6] Costa L da F, Rodrigues F A , Travieso G and Villas Boas P R 2007 Characterization of complex networks: A survey of measurements Advances in Physics 56(1) 167–242 doi:10.1080/00018730601170527

[7] De Domenico M, Porter M A, Arenas A 2015 MuxViz: a tool for multilayer analysis and visualization of networks Journal of Complex Networks 3 159-176 doi:10.1093/comnet/cnu038
[8] Lacasa L, Luque B, Ballesteros F, Luque J, and Nuño J C 2008 From time series to complex networks: The visibility graph Proceedings of the National Academy of Sciences 105(13) 4972–75 doi:10.1073/pnas.0709247105
[9] Melo D, Fadigas I and de Barros Pereira H B 2017 Categorisation of polyphonic musical signals by using modularity community detection in audio-associated visibility network Appl Netw Sci 2 32 doi:10.1007/s41109-017-0052-1
[10] Yela F, Thalmann F, Nicosia V, Stowell D and Sandler M 2019 Efficient On-line Computation of Visibility Graphs Delia. arXiv: 1905.03204v1 [cs.DS]
[11] Portnyagin D G, Sebyakin A G, Kuular E K, Trufanov A I, Berestneva O G and Tikhomirov A A 2018 Software used to identify signs of video editing. Proc. V Int. Scientific Conf. "Information Technologies in Science, Management, Social Sphere and Medicine" Part 2, ed Berestneva O G, Mitsel A A, Spitsyn V V and Gladkova T A (Tomsk: Publishing House of Tomsk Polytechnic University) 138-141
[12] Kuular E, Tikhomirov A and Trufanov A 2018 Two-component network model in voice identification technologies Security of Information technologies 1 81-89 doi: 10.26583/bit.2018.1.08
[13] Liu X, Tse C K and Small M 2009 Composing music with complex networks Int. Conf. on Complex Sciences: Theory and Applications, (COMPLEX2009) Shanghai 2196-2205
[14] Yang Q, Gao Q and Fan R 2011 A transformation from a singing voice to complex network using correlation coefficients of audio signals 2nd Int. Conf. on Intelligent Control and Information Processing 929-30
[15] Hurtado-Jaramillo J S, Guarin D L and Orozco A 2012 Complex networks: Application to pathology detection in voice signals Annual International Conference of the IEEE Engineering in Medicine and Biology Society 4229-32 doi: 10.1109/EMBC.2012.6346900
[16] Ferretti S 2017 On the complex network structure of musical pieces: analysis of some use cases from different music genres Multimedia Tools and Applications 77(13) 16003–29 doi: 10.1007/s11402-017-5175-y
[17] Park D, Bae A, Schich M, and Park J 2015 Topology and evolution of the network of western classical music composers EPJ Data Sci 4(1) 1 doi: 10.1140/epjds/s13688-015-0039-z
[18] Lacasa L, Nicosia V and Latora V 2015 Network structure of multivariate time series Scientific Reports 5(1) 1 doi:10.1038/srep15508
[19] Eckmann J P, Kamphorst S O and Ruelle D 1987 Recurrence Plots of Dynamical Systems Europhysics Letters 5(9) 973–977 doi:10.1209/0295-5075/4/9/004
[20] Marwan N, Donges J F, Zou Y, Donner R V and Kurths J 2009 Complex Network Approach for Recurrence Analysis of Time Series Physics Letters A 373 4246-54 doi:10.1016/j.physleta.2009.09.042
[21] Subramaniyam N P and Hyttinen J 2014 Characterization of dynamical systems under noise using recurrence networks: Application to simulated and EEG data Physics Letters A 378 (2014) 3464–74 doi:10.1016/j.physleta.2014.10.005
[22] Zhang J and Small M 2006 Complex network from pseudoperiodic time series: Topology versus dynamics Physical Review Letters 96 238701 doi: 10.1103/PhysRevLett.96.238701
[23] Small M, Zhang J and Xu X 2009 Transforming Time Series into Complex Networks Complex Sciences 2078–89 doi: 10.1007/978-3-642-02469-6_84
[24] Sun X, Small M, Zhao Y and Xue X 2014 Characterizing system dynamics with a weighted and directed network constructed from time series data Chaos: An Interdisciplinary Journal of Nonlinear Science 24 024402 doi:10.1063/1.4868261
[25] Zhang L, Small M and Judd K 2015 Exactly scale-free scale-free networks. Physica A: Statistical Mechanics and its Applications 433 182-197 doi:10.1016/j.physa.2015.03.074
[26] Gao Z, Dang W, Mu C, Yang Y, Li S and Grebogi C 2018 A Novel Multiplex Network-based Sensor Information Fusion Model and Its Application to Industrial Multiphase Flow System IEEE Transactions on Industrial Informatics 14(9) 3982-88 doi:10.1109/TII.2017.2785384
[27] Zhang Z, Xu J and Zhou X 2019 Mapping time series into complex networks based on equal probability division AIP Advances 9 015017 doi:10.1063/1.5062590
[28] Lan X, Mo H, Chen S, Liu Q and Deng Y 2015 Fast transformation from time series to visibility graphs. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8), 083105. doi:10.1063/1.4927835

[29] The Open Graph Viz Platform (Release 0.9.2/26 September 2017) URL: https://github.com/gephi/gephi