Topological Methods for Detection and Analysis of Cluster Structure in Complex Multidimensional Systems

Denys Symonov[1,*,†], Oleksandr Palagin[1,†] and Bohdan Zaika[1,†]

[1] V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences (NAS) of Ukraine, Academician Glushkov Avenue 40, 03187, Kyiv, Ukraine

Abstract

The paper considers the use of topological analysis, in particular the method of persistent homology and Vietoris-Rips complexes, to study the structural organization of complex multidimensional systems. It is shown that traditional clustering methods, such as k-means, are limited in identifying only compact subgroups. In contrast, Topological Data Analysis (TDA) allows identifying multidimensional coalitions, global and local cycles, isolated subsystems and topological barriers that determine the stability and functional integrity of the system. The proposed approach formalizes the concepts of coalition and topological barrier through the analysis of persistence barcodes and diagrams, providing quantitative identification of critical structural invariants in technical and information networks. Based on the modeling of a 200-element network, the ability of persistent homology to identify stable components, cycles, and isolated fragments even in the presence of noise or structural changes is demonstrated. Comparative analysis with classical metrics and k-means clustering confirmed the advantages of TDA in detecting multidimensional topology and increased robustness of the cluster structure. These results demonstrate the potential of TDA for analyzing complex systems.

Keywords

persistent homology, topological data analysis, Vietoris–Rips complex, complex multidimensional systems, cluster analysis, network robustness

1. Introduction

In modern science and technology, the study of complex multidimensional systems is of central importance due to the growing complexity of economic [1], electromechanical [2], sociotechnical [3], and medical structures [4]. Such systems, regardless of the field of application, demonstrate a multi-level organization, the presence of stable subsystems, dynamic coalitions, redundant and isolated components, which significantly affects their stability, adaptability, controllability, and ability to self-restore [5-7]. Identification, formalization, and quantitative analysis of structural invariants, such as stable coalitions, multidimensional barriers, and critical integration points, is essential for understanding the functioning of complex systems, ensuring their reliability, predicting degradation scenarios, and designing effective architectures. As the structural and behavioural complexity of such systems continues to grow, there is an increasing need for analytical methods capable of capturing high-dimensional, nonlinear, and topologically rich features beyond the reach of traditional approaches [8].

Classical network, graph, and statistical approaches provide a wide range of tools for analyzing local and global characteristics of systems (power distribution, clustering, centrality, modularity, etc.). However, these methods have fundamental limitations in the case of multidimensional

ICST-2025: Information Control Systems & Technologies, September 24-26, 2025, Odesa, Ukraine
* Corresponding author.
† These authors contributed equally.
denys.symonov@gmail.com (D. Symonov); palagin_a@ukr.net (O. Palagin); zaikabohdan5@gmail.com (B. Zaika)
0000-0002-6648-4736 (D. Symonov); 0000-0003-3223-1391 (O. Palagin); 0009-0001-9567-8361 (B. Zaika)

structures: they do not allow identifying higher-order topological patterns, such as global cycles, stable multi-level coalitions, complex redundant or autonomous subsystems that are not reducible to simple groups or nodes with increased connectivity [9-11]. In complex technical, information, or hybrid systems, there is a need for methods capable of detecting and quantifying multidimensional topological invariants that are critical to the integrity and functional stability of the entire structure.

Topological data analysis (TDA) offers a fundamentally new approach to the study of complex systems based on the analysis of persistent topological features at different scales. With the help of Vietoris-Rips complexes, persistence barcodes, and persistence diagrams, this approach allows us to identify not only individual clusters or isolated components, but also to detect multidimensional coalitions, global and local cycles, isolated "islands," and critical topological barriers. It is these structural elements that determine the system's resilience to disturbances, its ability to self-restore, functional integrity, and the presence of critical points, the destruction of which can lead to loss of control or fragmentation. Persistent homology, as a key TDA technique, is applied in this study through standard computational tools (Ripser, Gudhi), supplemented by custom scripts developed by authors to preprocess data and visualize persistent features relevant to the structural analysis.

This paper focuses on the use of persistent homology to formalize and quantify key coalitions and subsystems in complex multidimensional structures. The problem is formulated, the corresponding mathematical model is constructed, simulations are performed, and the interpretation of the obtained topological invariants in terms of functional stability and structural hierarchy of the system is proposed. A comparative analysis with classical network metrics confirms the unique analytical capabilities of TDA, in particular its ability to identify patterns that remain invisible to traditional approaches.

2. Problem statement

In complex multidimensional systems - regardless of their physical, technical or informational nature - structural coalitions of components play a key role in ensuring collective functionality, resilience and adaptability. Such coalitions can include compact subsystems with close interaction, as well as isolated fragments or multi-level associations that form critical functional blocks of the system. Classical approaches do not allow to identify multidimensional patterns of interconnection, such as global cycles, stable isolated subsystems, and redundant configurations. The problem is to formalize, identify, and quantify such structural invariants. It is solved by using topological analysis (TDA), in particular persistent homology, which allows identifying and interpreting stable coalitions, topological barriers, and critical points of system integration based on Vietoris-Rips complexes and persistence diagrams.

3. The goal of the research

The aim of this paper is to substantiate and demonstrate the effectiveness of topological analysis, in particular the method of persistent homology, for detecting, classifying and quantifying stable coalitions, critical subsystems and multidimensional topological invariants in complex multidimensional systems. The main tasks are: formalization of the concepts of coalition and topological barrier in the context of complex networks; construction of an appropriate mathematical model based on Vietoris-Rips complexes; algorithmic implementation of the analysis of persistence barcodes and diagrams; numerical modeling of structural scenarios; comparison of topological characteristics with classical network metrics to assess the informational content and practical value of the chosen approach.

4. Theoretical foundations of topological analysis of complex systems

Complex multidimensional systems can be defined as a set of agents $V = \{v_1, v_2, \dots, v_n\}$, between which there are likely to be connections of different nature (physical, informational, social, economic or other interactions) that determine the collective behavior of the system and its evolution over time. Formally, such a system is modeled as a graph $G = (V, E, W)$, where V is the set of n vertices, E is the set of m edges, and $= \{\omega_{ij}\}, i = \overline{1, n}$ is the set of edge weights, where each ω_{ij} characterizes the connection between vertices v_i and v_j.

The parameter ϵ (filtering parameter) sets the weight threshold at which vertices and their subsets are merged into simplices; it thus controls the density of connections at which multidimensional simplices (triangles, tetrahedra, etc.) are formed, reflecting coalitions, groups, and higher-order structures in a complex multidimensional system. Varying ϵ induces a filtration of the complex, allowing us to trace the evolution of topological characteristics. To study the deep structural organization and analyze multidimensional interactions in complex systems, the construction of simplicial complexes is used. In particular, Vietoris-Rips complexes are widely used [12, 13], which are formed as sets of subsets $\sigma \subseteq V$, such that for all pairs $v_i, v_j \in \sigma$ the edge (v_i, v_j) exists in E and $\omega_{ij} \geq \epsilon$:

$$VR_\epsilon(G) = (\sigma \subseteq V: \forall v_i, v_j \in \sigma, \omega_{ij} \geq \epsilon). \quad (1)$$

The constructed Vietoris-Rips complexes serve directly as the foundational structure for computing persistent homology, allowing for the identification of topological invariants, such as stable coalitions, barriers, and isolated subgroups, thus achieving the goals set in this study.

Persistent homology is a central tool for the topological analysis of complex systems and allows us to identify stable topological features at different scales [14, 15]. For a given filtering $\{K_\epsilon\}_{\epsilon \in [a,b]}$ of simplicial complexes, we consider the appearance and disappearance of homology classes: connectivity components (H_0), cycles (H_1), cavities (H_2), etc.

A persistent homology is formally defined as a sequence of homology groups:

$$H_1(K_{\epsilon_1}) \to H_2(K_{\epsilon_2}) \to \cdots \to H_k(K_{\epsilon_m}), \quad (2)$$

where H_k is the k-th homology group, where mappings are induced by the inclusion of complexes as ϵ increases.

In the persistent homology framework, the structural evolution of a complex system as the filtration parameter t increases is encoded by a persistence diagram $D_k = \{(b_i, d_i)\}_{i=1}^{N_k}$, where N_k is the number of homology classes of order k in the diagram. Here, each pair (b_i, d_i) represents the birth time b_i and death time d_i of a topological feature (such as a connected component, cycle, or cavity) in the filtration. The interval $[b_i, d_i)$ is called the persistence interval and characterizes the lifetime of the corresponding feature. Thus, the persistence diagram D_k is a finite multiset of points in $\mathbb{R}^2$, one for each homology class of order k.

For quantitative and statistical analysis, the persistence landscape function $\lambda_k(t)$ is used. This function is defined for the filtration parameter $t \in [0, T]$, where T is the maximal value considered in the filtration (e.g., the largest connection threshold in the Vietoris–Rips complex). The persistence landscape $\lambda_k(t)$ encodes, at each scale t, the maximum "height" of the topological features and enables the use of statistical and machine learning methods for further data analysis:

$$\lambda_k(t) = \sup_i[\min\{t - b_i, d_i - t, 0\}]. \quad (3)$$

In addition, for further numerical analysis of persistence diagrams, the persistence images are used, which are vectorized representations formed by projecting topological invariants into a fixed lattice of the birth-death space and then smoothing them. Such a representation allows using classical machine learning methods (e.g., SVM, PCA, neural networks) to solve classification problems, identify structural patterns, and distinguish between complex multidimensional systems.

For the quantitative comparison of topological features extracted from different states or versions of the system, we consider pairs of persistence diagrams D_k and D'_k of the same homology order k, corresponding to the original and modified networks, respectively. The choice of metrics for determining the relationships between agents is critical for building correct simplicial complexes. For undirected networks, Euclidean ($\ell_2 - \mathrm{norm}$), Manhattan ($\ell_1 - \mathrm{norm}$), cosine, or other distances in the feature space of agents are often used. For further comparison of topological structures, in particular persistence diagrams, specialized stability metrics such as bottleneck distance d_B and Wasserstein distance $d_{W,p}$ $[16]$ are used, which take into account differences in the position and duration of homologous classes:

$$d_B\left(D_k, D_k^{'}\right) = \mathrm{inf}_{\gamma}\mathrm{sup}_{x\in D_k}\|x-\gamma(x)\|_{\infty}, \tag{4}$$

$$d_{W,p}\left(D_k, D_k^{'}\right) = \left(\mathrm{inf}_{\gamma}\sum_{x\in D_k}\|x-\gamma(x)\|_{\infty}^{p}\right)^{\frac{1}{p}}, \tag{5}$$

where γ is the bijection between the points of the diagrams D_k and D'_k; $\|\cdot\|_{\infty}^{p}$ is the Chebyshev norm.

The correct choice of metrics and filtering parameters has a critical impact on the interpretation of topological structures and their stability in a multivariate system model. An unsuccessful setting can lead to the loss of significant patterns or, conversely, to the detection of artifacts caused by noise or excessive data complexity. Therefore, the stage of selecting metrics, filtering threshold ϵ, and scale of analysis is an integral part of valid topological modeling of complex multidimensional interactions. The optimal filtering parameter ϵ and distance metrics are typically selected empirically based on the stability and interpretability of persistent homology results, considering criteria such as persistence intervals stability and robustness against noise. In this study, these parameters were chosen iteratively, assessing multiple scenarios for best capturing structural invariants.

5. Topological indicators of the complex systems structure

5.1. Homology groups as indicators of structural features

Homology groups are fundamental algebraic objects that describe the topological structure of multidimensional systems. Formally, for a simplicial complex K, a homology group $H_k(K)$ is defined as a factor group:

$$H_k(K) = \frac{Z_k(K)}{B_k(K)}, \tag{6}$$

where $Z_k(K)$ is the group of k-cycles (k-chains with zero boundary); $B_k(K)$ is the group of k-boundaries (boundaries of (k+1)-chains).

Accordingly, the elements of $H_k(K)$ reflect:

- H_0: components of connectivity (clustering);
- H_1: independent cycles (barriers, “rings”);
- H_2: cavities (“empty” regions of higher dimension), etc.

A quantitative characteristic of homology groups is the Betti numbers (β_k), which are defined as the ranks of the corresponding homology groups:

$$\beta_k = \mathrm{rank}\, H_k(K), \tag{7}$$

where β_0 is the number of independent connectivity components, β_1 is the number of independent cycles, β_2 is the number of two-dimensional cavities, etc.

The Betti numbers $\beta_k(\epsilon)$ for each order k are defined as the ranks of the corresponding homology groups for the complex at the current value of the filtration parameter ϵ.

In the problems of dynamic system analysis, the filtering of complexes $\{K_\epsilon\}$ at a variable threshold ϵ, which produces persistent homologies, is considered. Each homology of class is characterized by the existence interval, which is defined by the persistence diagram for the k-th order $D_k = \{(b_i, d_i)\}_{i=1}^{k}$.

Persistence intervals are powerful indicators of structural features:

- long intervals (large $d_i - b_i$) indicate stable topological features (e.g., stable coalitions or barriers in social networks);
- short intervals correspond to "noise" or local changes in structure.

The invariants allow us to formalize the topological complexity of the system, compare it with other structures, and identify significant deviations or stable patterns at different scales. The following numerical invariants are used to quantitatively analyze the dynamics and compare the structures of different multidimensional systems:

- maximum value of persistence $\mu_k^{max} = \max_{(b_i,d_i)\in D_k}(d_i - b_i)$;
- average value of persistence $\mu_k^{mean} = \frac{1}{k}\sum_{i=1}^{k}(d_i - b_i)$;
- number of long-lived classes $n_k^{long} = |\{(b_i, d_i) \in D_k: (d_i - b_i) > \tau\}|$, where τ is a given stability threshold.

In addition, changes in Betti-numbers at different scales ϵ (Betti curve) are analyzed, which allows tracking structural transitions, the emergence and disappearance of coalitions or barriers in the network dynamics.

5.2. Classification of stable pattern types

Persistent homology, combined with numerical topological invariants such as Betti numbers, provides a formalized approach to detecting and classifying persistent patterns in complex multidimensional systems. The quantitative representation of persistence diagrams, $\beta_k(\epsilon)$ functions, and statistical characteristics of persistence intervals allows both interpreting persistent structures and automating their detection using machine learning methods. Within this classification scheme, three fundamentally different types of stable topological patterns can be can distinguished: coalitions, barriers, and isolated subgroups.

Coalitions (connectivity components, H_0) are formally identified as connected components of a graph or simplicial complex at a fixed filtering level ϵ. Here, coalitions specifically refer to connected components identified at a fixed filtering level ϵ. The introduction of the filtering parameter ϵ enables the identification of stable (persistent) coalitions and their changes as ϵ varies, distinguishing the approach clearly from classical definitions. The zero-order Betty number (β_0) is equal to the number of independent coalitions in the system:

$$\beta_0(\epsilon) = |\pi_0(K_\epsilon)|, \quad (8)$$

where $\pi_0(K_\epsilon)$ is the set of connectivity components of the space K_ϵ, and vertical dashes indicate its quantity (cardinality).

The dynamics of the number of connectivity components $\beta_0(\epsilon)$ during the filtering process allows tracing the processes of coalition (group unification) and fragmentation (group disintegration). Particular attention is drawn to the long-term coalitions that correspond to those components whose persistence intervals (b, d) are significantly higher than the average level, in other words $(d-b) \gg \langle d-b \rangle$, where $\langle d-b \rangle$ denotes the average persistence interval length, i.e., the mean lifetime of topological features in the persistence diagram

Barriers (cycles, H_1) are interpreted as one-dimensional topological cycles in the simplicial complex. They indicate the presence of structures that prevent the complete unification of subgroups or create isolation effects. The first-order Betti number $\beta_1(\epsilon) = \mathrm{rank}\, H_1(K_\epsilon)$ determines the number of independent barriers. The long-lived intervals (b, d) in the persistence diagram D_1 reflect persistent social or informational barriers that persist over a wide range of linkage parameters.

Isolated subgroups appear as components of connectivity with short persistence intervals. They appear at small ϵ and quickly disappear as the threshold increases. Quantitatively, such subgroups can be identified through the statistics of short intervals in the persistence diagram D_0:

$$n_0^{\mathrm{short}} = |\{(b_i, d_i) \in D_0 : (d_i - b_i) < \tau\}|. \tag{9}$$

If, for a given order k, there are no intervals satisfying the short-lived condition, then the corresponding set D_k^{short} is empty, and the number of such classes equals zero.

The classification of topological pattern types is based on quantitative analysis of persistence diagrams, statistical characteristics of persistence intervals (length, density, distribution), and the use of machine learning algorithms, where persistence images serve as vectorized features for automated recognition of structures in large multidimensional data.

5.3. Phase transitions in the topology of multidimensional systems

The analysis of phase transitions and structural transformations in complex multidimensional systems is a fundamental task, since such transitions are often accompanied by qualitative changes in the topology of connections, coalitions, distribution of influence, and information flows.

Since a multidimensional system is modeled as a simplicial complex $\{K_\epsilon\}$ with a filtering parameter ϵ (for example, a threshold of the connection strength or similarity). A phase transition is defined as a region of values of ϵ^* in which the topological invariants undergo a sharp change:

$$\exists \epsilon^* : \lim_{\delta \to 0} |\beta_k(\epsilon^* + \delta) - \beta_k(\epsilon^* - \delta)| \gg 0, \tag{10}$$

or

$$\left| \frac{d\beta_k}{d\epsilon} \right|_{\epsilon = \epsilon^*} \gg 1. \tag{11}$$

Typical examples of topological phase transitions are the following scenarios:

- With an increase in the filtering threshold ϵ, there is a transition from a set of isolated components ($\beta_0(\epsilon)$ is of great importance) to the formation of a giant connectivity component (a sharp decrease in $\beta_0(\epsilon)$), which reflects the integration of most elements of the system into a single structure.
- The appearance or disappearance of long-lived cycles ($\beta_1(\epsilon)$), which signals the formation or destruction of stable topological barriers or cyclic structures.

The main metrics for analyzing the trajectories of persistence landscapes are the discrete L^p-distance between the k-th landscapes at two filtration levels t_1 and t_2 (denoted as $d_{L^p}\left(\lambda_k^{(t_1)}, \lambda_k^{(t_2)}\right)$), the rate of the landscape change $V_k^{(t)}$, which acts as an indicator of the dynamic process activity, and the integral characteristic of changes over the entire period of time S_k:

$$d_{L^p}\left(\lambda_k^{(t_1)}, \lambda_k^{(t_2)}\right) = \left(\sum_m \int \left| \lambda_k^{m,(t_1)}(s) - \lambda_k^{m,(t_2)}(s) \right|^p ds \right)^{\frac{1}{p}}, \tag{12}$$

$$V_k^{(t)} = \left\| \lambda_k^{(t_2)} - \lambda_k^{(t_1)} \right\|_{L^2}, \quad (13)$$

$$S_k = \sum_{t=1}^{N-1} V_k^{(t)}. \quad (14)$$

Interpretation of trajectories as indicators of dynamics:

- Stable phases are characterized by low values of $V_k^{(t)}$, which indicates the preservation of the topological structure of the network.
- Phase transitions and anomalies are manifested as sharp peaks in the sequence $V_k^{(t)}$, which correspond to significant structural transformations.
- The path of a landscape in a multidimensional functional space can be subjected to clustering, component analysis, or principal component decomposition to identify the main stages of evolution.

Thus, phase transitions in multidimensional systems are manifested as abrupt changes in topological invariants, and metrics in the space of persistence landscapes (in particular, $V_k^{(t)}$, d_{L^p} and S_k) allow us to quantitatively record transformations, tracking the moments of structural reorganization and identifying dynamic phases of system development.

6. Theoretical foundations of topological analysis of complex systems

Verification of topological modeling results in complex multidimensional systems involves three key components: comparison with classical network metrics, analysis of the stability of invariants to data variations, and assessment of statistical significance. Although the paper mainly deals with weighted graphs, classical (unweighted) adjacency matrices are briefly mentioned here for completeness and better understanding of transition to weighted scenarios. In this context, the formalization of appropriate verification criteria that provide an objective assessment of the reliability, stability, and relevance of the obtained topological characteristics of the system is of particular importance.

The clustering coefficient is used to assess the extent to which nodes in the network tend to form local clusters or clustered groups [17]. It is calculated using the following formula:

$$C = \frac{1}{n} \sum_{i=1}^{n} \frac{2e_i}{k_i(k_i - 1)}, \quad (15)$$

where e_i is the number of edges between the neighbors of vertex v_i, k_i is its degree, $k_i \geq 2$.

If the graph G=(V,E,W) is weighted, then the weighted clustering coefficient is used:

$$C^{\omega} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{k_i(k_i - 1)} \sum_{\substack{j,h \\ j \neq h}} \left(\tilde{\omega}_{ij} \tilde{\omega}_{ih} \tilde{\omega}_{jh} \right)^{\frac{1}{3}}, \quad (16)$$

where $\tilde{\omega}_{ij}\tilde{\omega}_{ih}\tilde{\omega}_{jh}$ are the corresponding normalized weights, $\tilde{\omega}_{ij} = \frac{\omega_{ij}}{\max_{(i,j)\in E} \omega_{ij}}$; ω_{ij} is the weight of the edge between vertices *i* and *j* (element of the adjacency weight matrix).

Modularity Q measures how much denser the connections within clusters are than expected in a random model with the same degree distribution. For a graph $G = (V, E)$, which is divided into clusters, the calculation of Q can be performed using the formula:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \quad (17)$$

where A_{ij} is an element of the adjacency matrix, $A_{ij} = \begin{cases} 1, & \text{if } (i,j) \in E \\ 0, & \text{if } (i,j) \notin E \end{cases}$; k_i is the degree of vertex *i* ($k_i = \sum_j A_{ij}$); *m* is the total number of edges in the graph, $m = \frac{1}{2}\sum_{i,j} A_{ij}$; c_i is the cluster to which vertex *i* belongs; $\delta(c_i, c_j)$ is the delta function, $\delta(c_i, c_j) = \begin{cases} 1, & \text{if } c_i = c_j \\ 0, & \text{if } c_i \neq c_j \end{cases}$.

For a weighted graph $G = (V, E, W)$, the modularity Q can be calculated as:

$$Q = \frac{1}{2m}\sum_{i,j}\left[\omega_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j), \tag{18}$$

where k_i is the weighted degree of vertex *i*, $k_i = \sum_j \omega_{ij}$; *m* is the total sum of the weights of all edges in the graph, $m = \frac{1}{2}\sum_{i,j} \omega_{ij}$; c_i is the number of the community (cluster) to which vertex *i* belongs.

The correlation between the number of long-lived cycles β_1 and C (or Q) is analyzed by using the Pearson's coefficient r:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2\right)^{\frac{1}{2}}}, \tag{19}$$

where x_i and y_i denote the values of two different characteristics for the same network element (e.g., $x_i = \beta_1(i)$, $y_i = C(i)$, and the specific choice of characteristics is detailed in the corresponding text or figures).

Let $D(A)$ denote the set of all (b_i, d_i) points in the persistence diagram constructed for the adjacency matrix *A*. To test the stability of the invariants, the bottleneck distance (4) between the persistence diagrams before and after the data variations is considered:

$$d_B(D(A), D(A')) = \inf_\gamma \sup_{x \in D(A) \cup \Delta} \|x - \gamma(x)\|_\infty, \tag{20}$$

where $A = (a_{ij})$ is the original adjacency matrix, a_{ij} represents the connection between vertices v_i and v_j; A' is a modified version of the initial adjacency matrix *A*, which is obtained as a result of making changes to the network to test the stability of topological invariants; $D(A), D(A') \subset \mathbb{R}^2$ are the corresponding persistence diagrams considered as manifolds; $\Delta = \{(t,t) | t \in \mathbb{R}\}$ is the main diagonal (added with infinite multiplicity); γ is the bijection between the diagrams $D(A) \cup \Delta$ and $D(A') \cup \Delta$; $\|\cdot\|_\infty$ is the L_∞-norm.

The bootstrap method with *N* replications is applied to determine statistical significance. Let $r_k = \sum_{i=1}^{N} I\{\beta_k^{\mathrm{rand},i} \geq \beta_k^{\mathrm{obs}}\}$ be the number of repetitions in which the statistic was not less than the observed one, and $I\{\cdot\}$ is the indicator function. The Betti numbers $\beta_k(\epsilon)$ for each order *k* are defined as the number of independent homology classes of order *k* in the Vietoris-Rips complex constructed at the current value of the filtration parameter ϵ. For statistical significance estimation, these quantities are calculated separately for each order, so the number of repetitions *r* and the probability *p* should also be indexed as r_k, p_k respectively. Then, the estimates of the probability of an event with small samples with Laplace correction ("add-one" smoothing) are calculated by the formula:

$$p_k = \frac{r_k + 1}{N + 1}. \tag{21}$$

The application of the Laplace correction prevents $p = 0$ at a finite N.

Modern approaches suggest using additional invariants, such as topological entropy and diversity persistence diagrams, which reflect the diversity and unevenness of persistent structures:

$$E(D_k) = -\sum_{i=1}^{n} p_{k,i} \log p_{k,i}\,, p_{k,i} = \frac{d_i - b_i}{\sum_j (d_j - b_j)}, \quad (22)$$

where $p_{k,i}$ is the normalized length of the *i*-th persistence interval for homology of order *k*, and *n* is the total number of such intervals for the chosen order.

High entropy and diversity usually correspond to complex but stable structures, while their decrease signals the loss of pattern diversity under the influence of perturbations.

To sum up, the combination of various verification criteria (from comparison with classical network characteristics to testing the stability of invariants and assessing statistical significance) provides a comprehensive and objective verification of the reliability, informativeness, and scientific correctness of topological modeling of complex multidimensional systems.

7. Topological data analysis algorithms and software tools

In modern topological data analysis (TDA), specialized algorithms and software tools play a key role in efficiently computing persistent homologies even for complex multidimensional systems. The most well-known libraries are Ripser, Gudhi, and Dionysus. Ripser is focused on the fast computation of persistent Vietoris-Rips homology complexes using optimized data storage structures and boundary matrix reduction. Formally, the homology group H_k is computed by filtering the complexes $K_{\epsilon_0} \subseteq K_{\epsilon_1} \subseteq \ldots \subseteq K_{\epsilon_m}$, where at each level we consider k-chains C_k and boundary operators $\partial_k: C_k \to C_{k-1}$. The persistence is determined by the intervals of appearance and disappearance of homology elements of the groups when ϵ changes. Gudhi is a flexible platform for processing various types of complexes (Vietoris-Rips, Alpha, Witness, etc.) and supports the visualization of persistence diagrams and barcodes. Dionysus provides an interface for Python and C++, allowing the integration of TDA analysis into complex network data processing pipelines. All of these libraries implement algorithms with computational complexity that grows exponentially with the increase in homology order *k* and complex dimension (number of vertices). For Vietoris-Rips complexes, the complexity is usually $O(n^{k+1})$, which imposes a limit on the size of the analyzed networks.

8. Practical examples and modeling

To demonstrate the capabilities of TDA analysis, let us consider an artificially simulated social network of a large organization. The network consists of 200 members organized into 10 departments (20 people in each). Most departments have strong ties within them (0.7–1.0), weak ties between departments (0–0.3), and include 3 leaders who maintain additional intensive contacts with other leaders and participants from various groups. Thus, the network structure features densely connected subgroups, isolated departments, and weak intergroup contacts, reflecting a complex hierarchy and topological heterogeneity. The main task is to identify stable coalitions, leadership roles, barriers to information exchange, and critical points of network transformation using persistent homology.

Figure 1 presents a fragment of the weighted adjacency matrix $W = [w_{ij}]_{i,j=1}^{200}$ of the simulated social network, where each value w_{ij} indicates the strength of the connection between nodes *i* and *j*. Values range from 0 (no connection) to 1 (strong connection), with intermediate values corresponding to weaker ties between groups (if $i = j$ then value is 0). The network modeled in the case study with a clear structural hierarchy and identified leaders serves as the basis for applying topological analysis. Below are the main results of modeling using the Vietoris-Rips complex and persistent homology to identify stable coalitions, barriers, and critical subgroups in the system.

i \ j	1	2	3	4	5	6	…	198	199	200
1	0.0	0.81236	0.98521	0.91959	0.93138	0.84813	…	0.0	0.0	0.0
2	0.81236	0.0	0.79098	0.74184	0.78764	0.85682	…	0.0	0.0	0.0
3	0.98521	0.79098	0.0	0.79759	0.84855	0.82826	…	0.0	0.0	0.0
4	0.91959	0.74184	0.79759	0.0	0.81660	0.78140		0.0	0.0	0.0
…	…	…	…	…	…	…	…	…	…	…
196	0.05865	0.0	0.0	0.0	0.0	0.0	…	0.95037	0.98064	0.85968
197	0.0	0.0	0.83892	0.0	0.0	0.0	…	0.97130	0.94522	0.92711
198	0.0	0.84788	0.97159	0.0	0.0	0.0	…	0.0	0.96951	0.92047
199	0.27508	0.19942	0.12354	0.0	0.0	0.0	…	0.96951	0.0	0.87039
200	0.0	0.0	0.0	0.0	0.0	0.0	…	0.92047	0.87039	0.0

Figure 1: Fragment of the weighted adjacency matrix of the modeled network.

Figure 2 shows the general structure of the modeled social network, illustrating the hierarchical organization, distribution of subgroups, and positions of leaders. The graph clearly identifies ten compact subgroups (clusters), each of which has a high internal density of connections, corresponding to departments with strong intragroup cohesion. The central core is formed by leaders (marked in red) who serve as intergroup "bridges", ensuring the integration of clusters into a single communication structure. Such a topology creates the preconditions for the emergence of characteristic patterns of persistent homology - long-lasting components, numerous local cycles, and a limited number of persistent barriers.

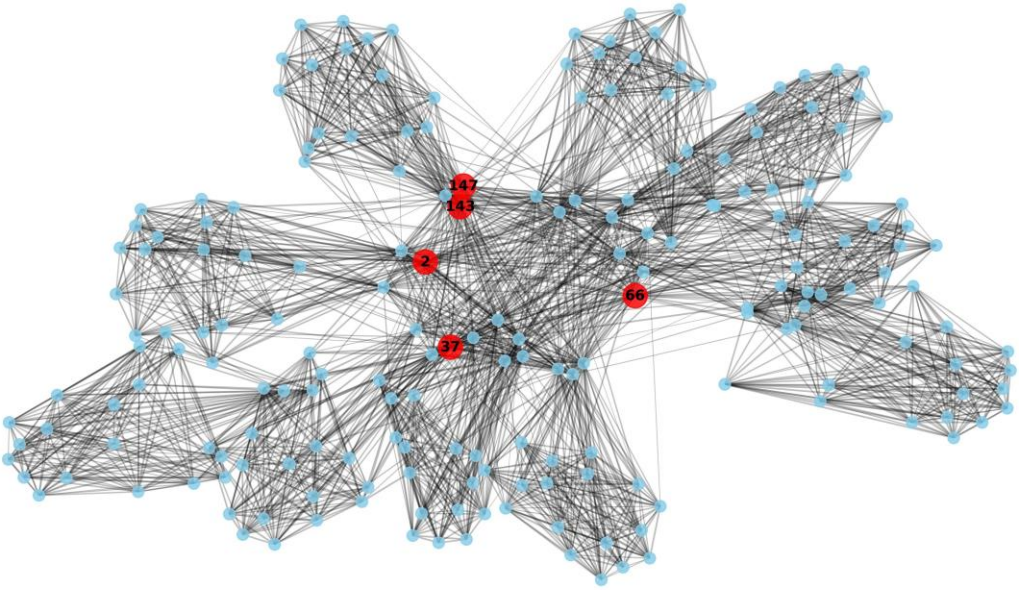

Figure 2: Global network structure: subgroups (blue) and leaders (red).

To compare the results of the topological and classical analysis, the data was clustered using the k-means method after multidimensional scaling (MDS), the results are shown in Figure 3. It can be seen that k-means forms ten compact clusters with well-defined boundaries in the space of the first two MDS components, but does not identify more complex topological structures, such as overlapping subgroups, stable cycles, or isolated components that are revealed by persistent homology. This confirms that classical clustering methods work well for distributed, almost convex groups, but are unable to detect multidimensional and hierarchical patterns inherent in complex systems, unlike the proposed method (Figure 2).

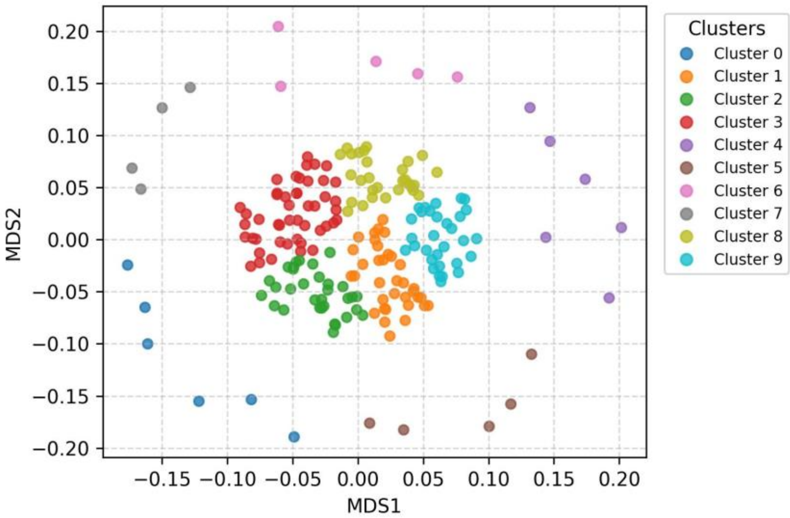

Figure 3: K-means clustering after MDS projection.

The dynamics of the topological invariants of the studied network presented in the form of a persistence barcode for the Vietoris-Rips complex is shown in Figure 4. At low values of the filtering threshold, there is a large number of short-lived connectivity components (β_0) that quickly merge into a single global component - this is shown by one long red line that persists until high threshold values. The blue barcode (β_1) illustrates multiple short-term cycles that quickly "fill in" as the threshold increases, as well as the presence of separate long-term cycles that correspond to stable local barriers in the network structure. The second group of cycles appears only at high thresholds, indicating isolated clusters with internal cohesion. This configuration of the persistence barcode reflects the presence of both rapidly integrated subgroups and autonomous, stable structures in the system, which is a key feature of complex social networks with a multi-level topology.

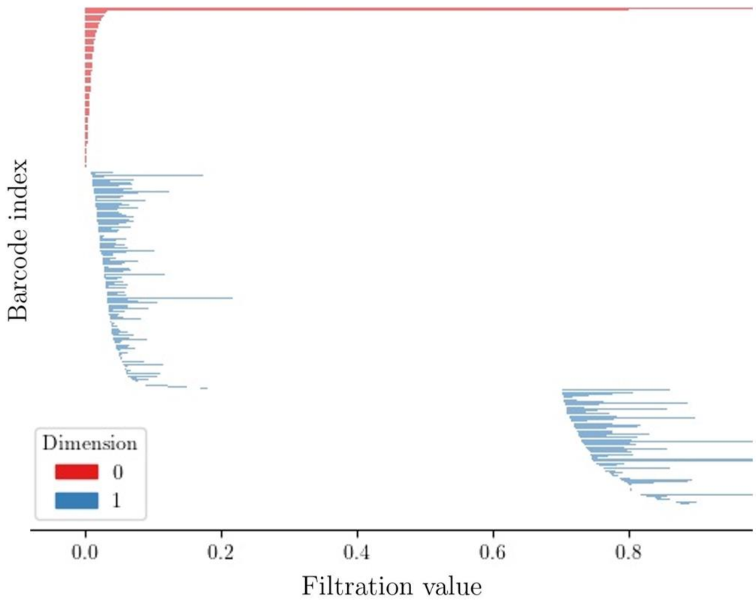

Figure 4: Persistence barcode for Vietoris-Rips complex of the modeled network.

To assess the robustness of the network topology, Figure 5 shows a comparison of persistence diagrams for the original, noisy, and modified (where leaders are attacked) versions of the network. The persistence diagrams for the Vietoris-Rips complex in all three scenarios demonstrate the preservation of key topological invariants: one dominant connectivity component (H_0) and two groups of cycles (H_1) - short-lived with small birth and death, and a group of stable cycles with large birth/death values. The introduction of random noise does not change the spatial configuration of the clusters of cycles, which indicates the stability of local coalitions to unstructured fluctuations. Even with the targeted removal of leaders, the main topological patterns are preserved, although some stable cycles disappear, which quantitatively illustrates the role of leaders as critical nodes for the integrity of the structure. Such robustness of the persistence diagram is a characteristic feature of the stability of a multidimensional network and confirms the high cohesion of the system core.

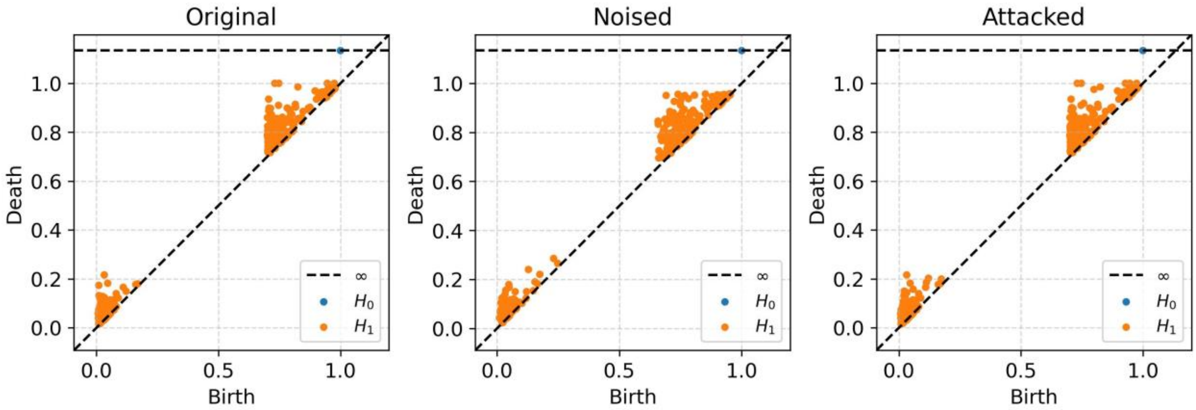

Figure 5: Persistence diagrams of Vietoris-Rips complex: original, noised and attacked network.

To characterize the structural heterogeneity of the network in detail, the Ego-network of the most and least active nodes was analyzed. Figure 6 illustrates the ego-networks of the most active (node 37) and the least active (node 134) elements within the analyzed structure. The ego-network of node 37 is characterized by a high density of internal and inter-cluster connections, forming a network hub topology with a minimal clustering coefficient and maximal betweenness centrality. This indicates the integrative role of this node in ensuring the global coherence of the network. In contrast, the ego-network of node 134 demonstrates localization, structural isolation, and a limited number of connections, reflecting its peripheral topological status and low impact on the system's dynamics. The observed topological contrast quantitatively confirms the stratified nature of the network organization and correlates with analytical findings obtained via both classical and topological methods for the study of complex systems.

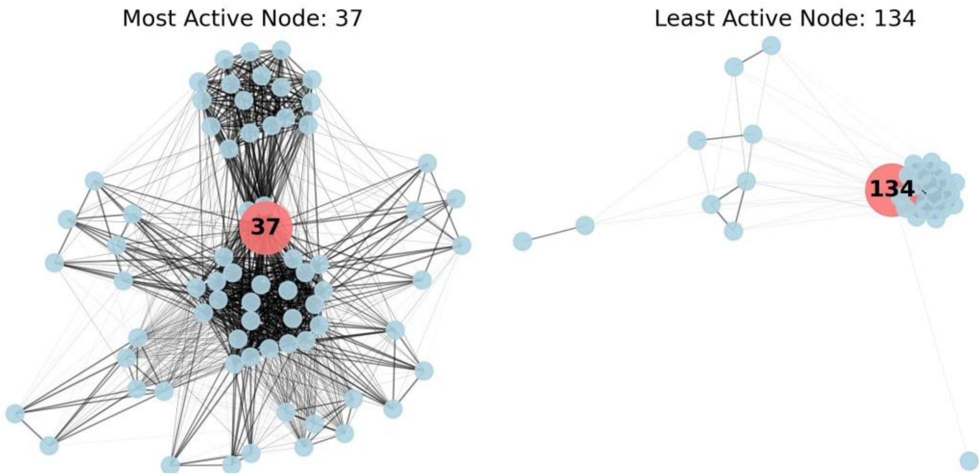

Figure 6: Ego-network of the most and least active nodes

To compare the topological and classical characteristics of nodes, the relationship between degree, clustering coefficient, and betweenness centrality is shown in Figure 7. There is a strong negative correlation between node degree and clustering coefficient: nodes with high degree have low clustering coefficient, which corresponds to the role of leader hubs that connect different subgroups with minimal internal connections. At the same time, there is a clear positive correlation between degree and betweenness centrality: nodes with the highest degree demonstrate the highest values of betweenness centrality, acting as critical structural intermediaries for intercluster integration of the network. Such distributions quantitatively confirm the hierarchical organization and functional differentiation of nodes in the modeled social system.

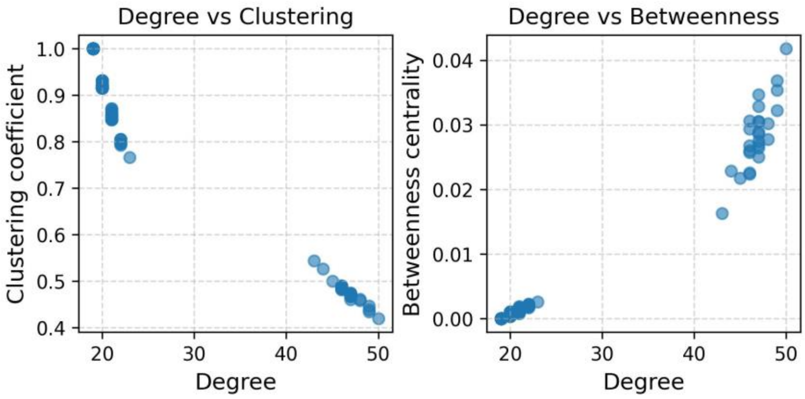

Figure 7: Relationship between clustering coefficient, betweenness centrality and node degree.

To quantitatively verify the visually observed relationships between the key network metrics, Spearman's rank correlation was calculated for 200 sample nodes. The analysis showed a very strong negative association between the clustering coefficient and the node degree ($\rho = -0.96$, $p < 0.001$), which confirms the tendency of high-degree nodes to lose local cohesion. At the same time, a very strong positive correlation was found between the node degree and the betweenness centrality ($\rho = 0.94$, $p < 0.001$), as well as a very strong negative correlation between the clustering coefficient and betweenness centrality ($\rho = -0.89$, $p < 0.001$). The obtained values remain statistically significant after the Holm correction for multiple comparisons, which indicates the extraordinary stability of the identified patterns.

The modeling results show that the topological analysis of a social network allows for the quantitative identification of stable coalitions, isolated subgroups, and critical leaders that ensure the global integration of the structure. Persistence barcodes and diagrams clearly reflect the hierarchical, clustered, and robust organization of the network, in which central hub nodes form stable connections between groups even in the face of perturbations or targeted attacks. Classical network metrics additionally emphasize the functional differentiation of the roles of participants and confirm the high cohesion of subgroups and the structural heterogeneity of the system. The proposed approach provides an in-depth interpretation of the multidimensional network structure and demonstrates the high sensitivity and reliability of TDA for identifying key topological patterns in complex multidimensional systems.

9. Possibilities and limitations of the topological approach

The topological approach to analyzing complex multidimensional systems is effective for identifying multidimensional group interactions, complex coalitions, isolated subgroups, and hidden barriers to information flow. TDA methods are particularly valuable for investigating collective dynamics, including the formation of stable associations, identification of marginalized or resilient groups, and analysis of influence centers' emergence and collapse. Persistent homology, through topological invariants $\beta_k(K_\epsilon)$, facilitates tracking coalitions' and barriers' transformations and identifying critical points of system fragmentation or integration.

However, practical applications of the topological approach encounter significant limitations. Firstly, computational complexity for Vietoris-Rips complexes typically scales as $O(n^{k+1})$, restricting analysis primarily to moderately sized networks or lower-order homologies. For large networks or higher-dimensional homologies, computational resources and runtime become prohibitive, necessitating approximation or simplification techniques. Secondly, results are sensitive to parameter selection-filtering thresholds, metrics, and noise levels-leading to potential misinterpretations or artifacts. This sensitivity mandates careful parameter tuning, robustness analysis, and validation procedures. Finally, current TDA implementations lack sufficient scalability and adaptability for heterogeneous, dynamic, or temporal data, highlighting the ongoing need for optimized algorithms and hybrid methodologies integrating TDA with classical statistical or machine learning tools.

10. Conclusions

In this paper, we demonstrate the effectiveness of topological analysis, in particular, persistence homology and Vietoris-Rips complexes, for the detection and quantitative interpretation of stable coalitions, barriers, and isolated subsystems in complex multidimensional systems. The modeling results showed that persistence diagrams and barcodes allow us to identify not only local clusters but also global topological patterns: the number of independent coalitions (β_0), stable cycles (β_1), and multidimensional group interactions that remain unchanged even with significant modifications of the structure. For example, the main coalition kernels and groups of long-lived cycles are preserved after adding noise or removing leaders, which confirms the robustness of the system. Comparative analysis with k-means clustering showed that classical methods capture only compact groups, while TDA allows identifying complex multidimensional barriers and hierarchies. Statistical tests confirmed the significance of the topological findings. In general, the proposed approach provides a high level of interpretability and objectivity for the analysis of critical structural invariants in complex systems.

Acknowledgements

The work was supported by the state budget research project "Develop methods for modelling the processes of targeted management of complex multi-component information systems for various purposes" (state registration number 0123U100754) of the V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences (NAS) of Ukraine.

Declaration on Generative AI

During the preparation of this work, the authors used DeepL in order to translate research notes and results from Ukrainian to English. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

[1] S. T. Shutters, Urban economic structures as multidimensional networks: A complex systems framework for analyzing economic development trajectories, Complexity 2024 1 (2024). doi:10.1155/2024/5521625.

[2] R. Wang, T. Li, Z. Gao, X. Yan, J. Wang, Z. Wang, J. Gao, A generative adversarial networks based methodology for imbalanced multidimensional time-series augmentation of complex electromechanical systems, Applied Soft Computing 153 (2024). doi:10.1016/j.asoc.2024.111301.

[3] F. Crestelo Moreno, V. Soto-López, D. Menéndez-Teleña, J. Roca-González, J. Suardíaz Muro, C. Roces, M. Paíno, I. Fernández, L. A. Díaz-Secades, Fatigue as a key human factor in complex sociotechnical systems: Vessel traffic services, Frontiers in Public Health 11 (2023). doi:10.3389/fpubh.2023.1160971.

[4] O. Palagin, D. Symonov, T. Semykopna, Mathematical modeling of the evolution of the rehabilitation process for patients with oncological diseases, in: V. Kazymyr et al. (Eds.), Mathematical Modeling and Simulation of Systems (MODS 2023), volume 1091 of Lecture

Notes in Networks and Systems, Springer, Cham, 2024, pp. 99–112. doi:10.1007/978-3-031-67348-1_8.
[5] X. Lin, Z. Shi, B. Wang, Stability of highly nonlinear neutral stochastic delay system with both stable and unstable subsystems, Int. J. Robust Nonlinear Control 34 (2024) 8355–8373. doi:10.1002/rnc.7394.
[6] F. Lezama, T. Pinto, R. Faia, Z. Vale, Exploiting flexibility of end-users at DSO-TSO levels acting in dynamic coalitions, in: Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2020, pp. 1–5. doi:10.1109/PESGM41954.2020.10041864.
[7] N. Sedaghat, M. Abouei Ardakan, G-mixed: A new strategy for redundant components in reliability optimization problems, Reliability Engineering & System Safety Safe. 216 (2021). doi:10.1016/j.ress.2021.107924.
[8] D. Symonov, O. Palagin, Y. Symonov, B. Zaika, Optimisation of training samples with KLE and mutual information, in: Proceedings of the Eighth International Workshop on Computer Modeling and Intelligent Systems, CMIS 2025, CEUR-WS.org, Zaporizhzhia, 3988, pp. 296–306. URL: https://ceur-ws.org/Vol-3988/paper23.pdf.
[9] A. N. Pisarchik, N. Peña Serrano, W. Escalante Puente de la Vega, R. Jaimes-Reátegui, Hypergraph analysis of functional brain connectivity during figurative attention, Applied Sciences 15 (2025). doi:10.3390/app15073833.
[10] M. Contisciani, F. Battiston, C. De Bacco, Inference of hyperedges and overlapping communities in hypergraphs, Nature Communications 13 (2022). doi:10.1038/s41467-022-34714-7.
[11] Q. F. Lotito, F. Musciotto, A. Montresor, F. Battiston, Higher-order motif analysis in hypergraphs, Commun. Phys. 5 (2022). doi:10.1038/s42005-022-00858-7.
[12] H. Adams, J. Bush, Ž. Virk, The connectivity of Vietoris–Rips complexes of spheres, J. Appl. Comput. Topol. 9 (2025). doi:10.1007/s41468-025-00214-y.
[13] M. S. Khan, A. Tiwari, Q. M. D. Lohani, Vietoris-Rips complex induced by intuitionistic fuzzy distance measure, in: Proceedings of the 2023 IEEE International Conference on Fuzzy Systems (FUZZ), IEEE, 2023, pp. 1–6. doi:10.1109/FUZZ52849.2023.10309679.
[14] A. T. Jafadideh, B. M. Asl, Topological analysis of brain dynamics in autism based on graph and persistent homology, Computers in Biology and Medicine 150 (2022). doi:10.1016/j.compbiomed.2022.106202.
[15] D. R. Carmody, R. B. Sowers, Topological analysis of traffic pace via persistent homology, Journal of Physics: Complexity 2 (2021). doi:10.1088/2632-072X/abc96a.
[16] R. U. Gobithaasan, K. D. Selvarajh, K. T. Miura, Clustering Datasaurus Dozen using bottleneck distance, Wasserstein distance (WD) and persistence landscapes, Journal of Advanced Research in Applied Sciences and Engineering Technology 38 (2024) 12–24. doi:10.37934/araset.38.1.1224.
[17] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442. doi:10.1038/30918.