

Hierarchical Spatio-temporal Visual Analysis of Cluster Evolution in Electrocardiography Data

Sugeerth Murugesan
Lawrence Berkeley National
Laboratory
smurugesan@lbl.gov

Max Dougherty
Lawrence Berkeley National
Laboratory
maxdougherty@lbl.gov

Kristofer Bouchard
Lawrence Berkeley National
Laboratory
KEBouchard@lbl.gov

Bernd Hamann
University of California, Davis
bhamann@ucdavis.edu

Edward Chang
University of California, San
Francisco
ChangEd@neurosurg.ucsf.edu

Gunther H. Weber
Lawrence Berkeley National
Laboratory
ghweber@lbl.gov

ABSTRACT

We present ECoG ClusterFlow, a novel interactive visual analysis tool for the exploration of high-resolution Electrocardiography (ECoG) data. Our system detects and visualizes dynamic high-level structures, such as communities, using the time-varying spatial connectivity network derived from the high-resolution ECoG data. ECoG ClusterFlow provides a multi-scale visualization of the spatio-temporal patterns underlying the time-varying communities using two views: 1) an overview summarizing the evolution of clusters over time and 2) a hierarchical glyph-based technique that uses data aggregation and small multiples techniques to visualize the propagation of clusters in their spatial domain. ECoG ClusterFlow makes it possible 1) to compare the spatio-temporal evolution patterns across various time intervals, 2) to compare the temporal information at varying levels of granularity, and 3) to investigate the evolution of spatial patterns without occluding the spatial context information. We present a case study for real epileptic seizure data aimed at evaluating the effectiveness of our approach.

Categories and Subject Descriptors

J.3 [Computer Applications]: Life and Medical Sciences—*Medical information systems*; H.4 [Information Systems]: Human factors—*Human information processing*; I.5.3 [Information storage and retrieval]: Information Search and Retrieval—*Clustering*

General Terms

Design, Human Factors

Keywords

Electrocardiography, Neuro-informatics, Dynamic Graphs, Graph Visualization

1. INTRODUCTION

The human brain is a highly connected, dynamic system comprising specialized brain regions that coordinate and interact in many complex ways for communication, producing intricate patterns of system behavior [17]. Analyzing these communication patterns can help us gain an understanding of the normal functioning of the brain, how we learn or age, and how neurological disorders develop or can be treated [17, 6]. Brain systems function across a large range of spatial and temporal scales. Investigating how the connectivity patterns vary across these different scales has provided new insights into how low-level signals cause global brain state transformations [20]. To support such analysis and capture these patterns comprehensively, data with high spatio-temporal resolution and low signal-to-noise ratio is needed.

Recent advances in invasive monitoring technologies such as electrocardiography (ECoG) have risen to this challenge by recording high-resolution electrical signals captured by electrodes placed directly on the cortical surface of the brain. The correlation of electrical activity between these electrodes yields a measure of functional connectivity between them. As the derived functional network changes over time, the topology and the attributes of the network vary as well, making it difficult to analyze and visualize the network.

Developments in graph theoretical methods have made it possible to simplify and characterize the data contained in the connectivity network. For example, through community detection methods, it has been determined that brain networks exhibit modular organization [16], i.e., they consist of clusters, subsets of regions having strong inter-modular connections and sparse inter-modular connections. These clusters represent specialized behavioral systems such as higher-order vision, or sensory-motor processing [25].

One way to explore how these behavioral systems interact when performing a task or are impaired due to neurological disorders is to study how the modules evolve over time [6]. This study involves identifying cluster evolution patterns such as: spatial distribution, or a combination of clusters; electrical activation or deactivation of a cluster; and the birth and death of clusters. In the case of epilepsy, for instance, visual analysis of the cluster data combined with the electrical activity can help differentiate normal and ictal (seizure) states of the brain. These patterns—when vali-

© 2016 Association for Computing Machinery. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of the United States government. As such, the United States Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

BCB'16, October 02-05, 2016, Seattle, WA, USA

© Copyright 2016 ACM. ISBN 978-1-4503-4225-4/16/10 \$15.00.

DOI: <http://dx.doi.org/10.1145/2975167.2985688>.

dated with statistical analysis—are crucial for a successful treatment of the identified epileptogenic zones.

The spatio-temporal patterns in time-varying clusters operate at different spatial and temporal scales. To capture and analyze these patterns, it is important that the temporal scale of the analysis matches the temporal scale of the patterns themselves [2]. For example, patterns such as spatial distribution or combination of clusters are best captured at a lower temporal scale while global transitions of brain states are captured at a higher temporal scale. Analyzing the patterns at varying granularity is crucial as appropriate scales for evaluation are not obvious *a priori* and a single optimal solution at a particular scale is unlikely to exist [2].

Existing approaches to visualize dynamic spatio-temporal clusters operate mostly at a single temporal scale and do not satisfactorily support the in-depth comparison and evaluation of the evolution patterns underlying the data. They mainly focus on visualizing such data by directly depicting all of the information through visual representations or using computational methods to reduce and summarize the visual data. While direct depiction methods suffer from scalability issues, data reduction methods ignore the low-level details that are important in explaining high-level patterns.

To support a comprehensive and detailed study of ECoG data, we present ECoG ClusterFlow (Fig. 1), an interactive system that supports the exploration, comparison and analysis of time-varying community evolution patterns at varying temporal granularity. Two major views of the system include: 1) an overview summarizing the overall changes in cluster evolution, where users explore salient dynamic patterns; and 2) a hierarchical glyph-based timeline visualization for exploring the dynamic spatial organizational changes of the clusters using data aggregation [9] and small multiples [36] methods.

These techniques allow users to gain insights at many levels of temporal granularity, exploring globally evolving patterns to observing small-scale spatial changes. In summary, our main contributions include:

- A hierarchical multi-scale approach to visualizing temporal modular changes in brain networks
- Unique glyph-based designs that explore spatial organizational changes of the dynamic cluster data

Furthermore, the specific design goals and capabilities of our system were articulated in close collaboration with the neuroscientists and the neurosurgeon on our team, ensuring that our prototype improves the overall data exploration process. Our system was repeatedly evaluated and tested by the users, making possible the development of analysis modules that help gain new insights into the data. We present two case studies using synthetic and epileptic seizure datasets to demonstrate the usefulness of our system.

2. RELATED WORK

Work related to ours falls into visualization of communities for dynamic graphs; and visualization of spatio-temporal data:

2.1 Communities for Dynamic Graphs

When exploring communities in dynamic graphs, existing techniques primarily use animation (time-to-time mapping) or static timeline-based (time-to-space mapping) visualization methods to depict modular changes over time.

In **animations**, the community structure of the network is shown by color-coding the nodes or partitioning the draw-

ing space into sections [8] or nested blocks [29] (if the data is hierarchical). Due to their reliance on short-term memory, animations increase the cognitive load during analysis [5]. A way to mitigate this problem is to maintain the ‘mental map’ of the layout by minimizing node movement in the animation [26]. An alternate approach to decrease the cognitive load is that place multiple graph representations along a timeline using small multiples [36]. However, this multi-view approach leaves the user with the manual task of assimilating and identifying changes.

To address this problem, several approaches utilize **timeline-based representations** [31, 28], visualizing only the evolution of clusters over time. In a timeline view, each segment along the axis perpendicular (to the timeline) represents a cluster identified at that particular timestep. The links between two axes represent the changes in the cluster affiliation of the nodes. The arbitrary ordering of the nodes in the vertical axis may increase link crossings between axes, inhibiting easy comprehension of the evolution patterns. To address this issue, Reda *et al.* and Sallabury *et al.* [28, 32] employ sorting techniques to place active and stable communities at the top of the vertical axis.

To further support the comprehension of transitions between communities, alluvial diagrams [30] model the links between clusters in different vertical axes as split-merge ribbons [31, 37]. This approach enhances the visual traceability of important cluster evolution patterns.

Most published methods either visualize the evolution of clusters or their corresponding graph topology. Reda *et al.* [27] visualize the spatial context and the evolution of the time-varying clusters by linking a space-time cube (3D spatial representation) with a timeline-based representation.

In contrast, our work provides the spatial context for dynamic evolution patterns in 2D. Furthermore, our technique matches clusters through an overall match algorithm, ensuring intuitive identification of dynamic community patterns.

2.2 Spatio-Temporal Data

Previous visual analysis methods for spatio-temporal data utilize either integrated or separated views [4].

Integrated views visualize spatial and temporal data in one view. Superimposing temporal graph data onto a spatial view [3] and visualizing a 3D space-time cube timeline over a 2D spatial view are two examples of integrated views [12]. Another hybrid 2.5D approach proposed by Tominski *et al.* [35] displays temporal information on top of a 2D spatial layout. However, for a large number of timesteps or data points, these views can easily become cluttered and occluded.

Separated views overcome visual clutter by using dedicated views to present different aspects of spatio-temporal data. Plug *et al.* [23] link data in spatial and temporal domains by using small multiples of maps, superimposing a subset of temporal data on each of the spatial maps. Jern *et al.* [18] utilize color to link spatial and temporal data. Other methods [22] for static data use interaction techniques to link data in both domains, requiring substantial and concentrated eye movements for visual analysis.

Visual glyph designs that aggregate spatio-temporal attributes not only reduce the size of the represented data but also enable intuitive comparison of temporal data. *Glidgets* [19] depicts temporal changes by segmenting glyphs into time slices, enabling the comparison of attributes over time. Nan Cao *et al.* [7] and Erbacher *et al.* [10] aggregate tem-

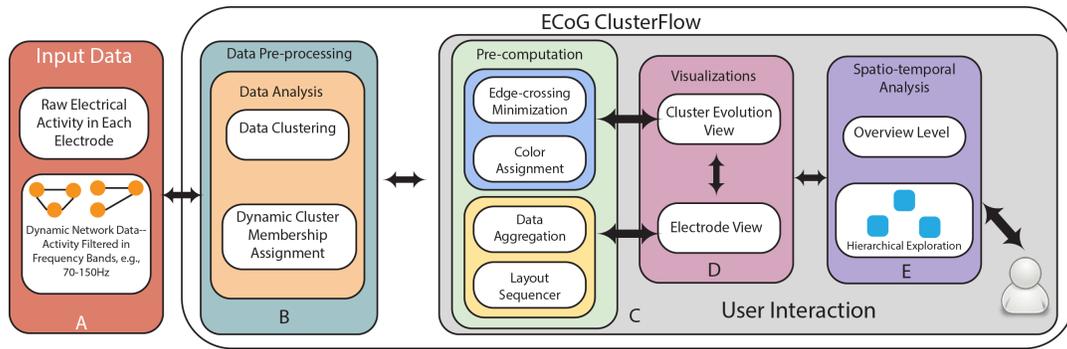


Figure 1: Overview of the ECoG ClusterFlow pipeline. A. Raw electrical signals are statistically analyzed to derive the dynamic network data. B. The data pre-processing step identifies and links cluster across timesteps. C. Main modules of the visualization system. D. Users can investigate patterns in two major visualization views. E. Users can perform various types of spatio-temporal analysis based on these views.

poral data to summarize the entire dataset with the overall goal of detecting anomalous behavior in the network. ECoG ClusterFlow utilizes some of the aforementioned concepts to provide unique glyph-based designs and visualization methods that show the overall modular changes of the network.

3. CLUSTER DETECTION

Our visualization methods are based on sequence of communities detected at each timestep. We call this sequence of communities **dynamic communities**. Given the graph at a particular timestep $G = \{N, E\}$, where N are the nodes that represent electrodes and E are the edges that represents the correlation between the electrodes, the community detection algorithm clusters the data into K non-overlapping and exhaustive communities. A K -cluster partition P , with communities C_i (where $1 \leq i, j \leq K$), is defined as $P = \{C_1, C_2, C_3, \dots, C_K\}$ such that $\sum_{k=1}^K C_k = G$ and $C_i \cap C_j = \emptyset, \forall i, j : i \neq j$.

Derivation of time-dependent clusters is an essential task in the analysis of time-varying brain network [33]. Two main approaches [11] are commonly used: 1) A two-stage approach where the communities are derived at each timestep and then tracked over time using different community tracking methods [13, 32]. 2) an evolutionary clustering approach that takes into account the graph topology and the clustering results from previous timesteps. However, based on the feedback from the neuroscientists on our team and other existing work [32, 33], we choose the two-stage clustering approach (described in detail in section 4.1.1) with consensus clustering [21] as our primary detection algorithm. This method produces a better quality of clustering results since each timestep is clustered locally (determining the correct number of clusters) [32], and combines the best outputs of multiple runs of the K-means clustering algorithm.

4. SYSTEM OVERVIEW

We developed ECoG ClusterFlow in close collaboration with neuroscientists (including neurosurgeons) to guide the design of our analysis framework and to ensure that it would be truly valuable as an exploratory tool.

Fig. 1 shows the pipeline of our system. The input to our system is: 1) the processed electrical signal data originating from each electrode in the ECoG grid and, 2) its

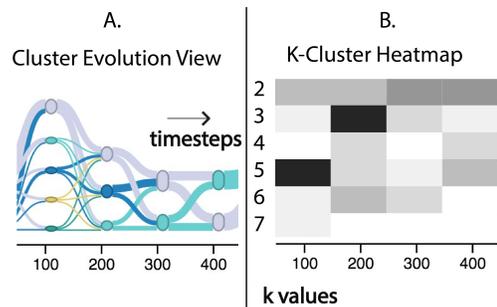


Figure 2: Evolution of clusters for four timesteps. A. The cluster evolution view shows clusters and transitions between them. The nodes have colors based on their cluster membership. B. The K-cluster heatmap on the bottom visualizes the likelihood of a range of K values that determine the final number of clusters for a particular timestep, for e.g. a clear maxima is evident for 100ms (5) and 200ms (3).

corresponding pairwise dynamic correlation network. The dynamic network data is pre-processed to derive dynamic clusters. Visualization methods, such as data aggregation, are applied to the cluster data in the pre-computation phase and final visualizations are generated.

Based on our conversations with domain experts and the network task taxonomy by Ahn et al. [1], we have identified the following domain questions of interest:

Identify temporal brain states(Q1) What activation patterns are consistent over a continuous period of time?

Identify transitions between brain states(Q2) Given the brain states, what patterns characterize their transition to another state?

Compare the evolution patterns associated with different brain states(Q3): What patterns underlie the brain states during normal versus diseased condition?

Assess changes in community membership(Q4) Given a spatial region of interest in the brain, how do the clusters belonging to these regions change over time?

These questions led us to establish the following system design goals:

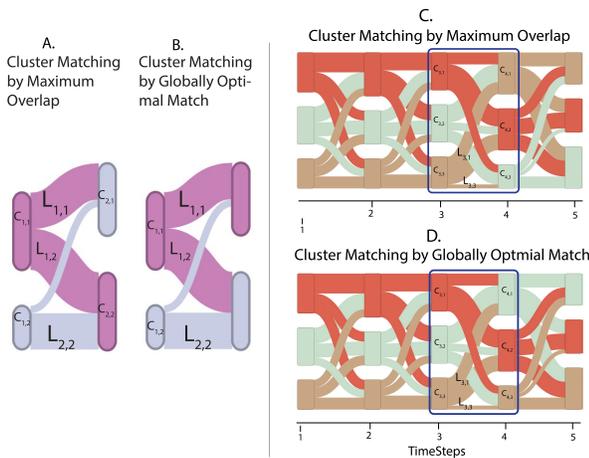


Figure 3: Comparison of tracking algorithms in artificial datasets. In image C the maximum overlap algorithm assign $C_{3,3}$ to $C_{4,1}$, while in Fig. D the globally optimal matching algorithm assigns $C_{3,3}$ to $C_{4,3}$, qualitatively making communities more visually traceable in image D. The scalar values for the links in image A and B are $L_{1,1} = 11$, $L_{1,2} = 11$, $L_{2,1} = 10$, $L_{2,2} = 2$.

Timeline-based visualizations(G1) Support views that display the time-varying cluster information on a static display to take advantage of the user’s visual perception instead of cognition (time-to-time mapping)

Multiple levels of detail and abstraction(G2) Support views that enable neuroscientists to explore the data at multiple levels of granularity for analysis

Holistic visualizations(G3) Support visual designs that combine multiple data attributes like cluster membership and its electrical activation

These goals are addressed in our system by two major views: the *Cluster Evolution View* and the *Electrode View*.

4.1 Cluster Evolution View

The cluster evolution view highlights the salient patterns of the cluster evolution including the emergence, death, contraction, expansion, merging and splitting of clusters (Q2, Q3, Q4). Through this view, analysts can compare and analyze modular signatures (cluster evolution patterns) over time and identify important time intervals and distinct brain states. The cluster evolution patterns are represented using a flow-based visualization [30, 37] (G1) (alluvial diagram), where the clusters metaphorically flow like a river with split/merge tributaries from left to right.

Formally, at each timestep t on the horizontal axis, rectangular blocks represent clusters $C_{t,i}$ where the height of each block represents the cluster’s size at that timestep. Flow-based transition links $L_{i,j}$, where i is the source community and j is the sink community, connect clusters to show changes in the community structure over time. We model these links as Bezier curves, to generate a continuous representation of the transition between successive communities [37]. Fig. 3 shows the evolution of dynamic clusters for five timesteps. Furthermore, to easily assess the community membership in dynamic clusters, we color communities us-

ing solid coloring, using N perceptually distinct colors from a qualitative colorbrewer [15] colormap.

4.1.1 Cluster Tracking

To support the two-stage cluster detection approach, correspondences between clusters in consecutive timesteps need to be determined. Based on the input from neuroscientists, we have investigated two approaches to compute this matching: (1) maximum overlap tracking and (2) computing the globally optimal match.

Maximum overlap tracking (Fig. 3A, C) is a greedy algorithm that iteratively matches the two clusters in consecutive timesteps that share the maximum number of electrodes. This process is repeated until no overlapping clusters remain. This approach may not always produce an intuitive correspondence between clusters. For example, in Fig. 3A and C, clusters $C_{1,2}$ and $C_{2,2}$ have maximum overlap (of 11 electrodes) and are paired in the first iteration. This only leaves $C_{2,1}$ as possible match for $C_{1,2}$ in the second iteration, even though the overlap between $C_{1,2}$ and $C_{2,1}$ is relatively small (only two electrodes).

To find the globally optimal assignment, our second approach picks the best overall match between all clusters in consecutive timesteps. We define a similarity measure

$$sim = \frac{|C_{t,i} \cap C_{t+1,j}|}{|C_{t,i} \cup C_{t+1,j}|}$$

between clusters C_i and C_j in consecutive timesteps t and $t+1$, similar to the approaches by Greene et al. [14] and Sallabury et al. [32]. Next, we compute a similarity matrix comprised of the pairwise similarity measures between all possible cluster combinations. To avoid matching of clusters with small overlap, we set to zero those similarity values that are below a threshold θ . To match clusters, we consider all possible cluster matchings between timesteps—by considering all possible permutations of clusters—and compute a global similarity value as the sum of the similarity values for all matched clusters. The overall best match is the permutation that maximizes global similarity. While considering all possible permutations is computationally expensive, we usually consider only upto a small number of clusters (approximately seven) per time step, keeping this approach tractable. Fig. 3D shows the best overall match for our example, matching clusters $C_{1,1}$ and $C_{2,1}$ as well as clusters $C_{1,2}$ and $C_{2,2}$, a more intuitive choice than the result obtained by maximum overlap tracking. An implementation of this approach for an artificial dataset with three dynamic communities is shown in Figure 3. The two approaches differ in the community results starting at timestep four.

4.1.2 Sorting and Ordering of Nodes

To enhance the visual traceability of the clusters, the node layout of the graph should ideally minimize edge crossings with optimal ordering of the nodes (clusters) at each vertical axis. To determine such an ordering, we must take all the timesteps into consideration. Several methods have been proposed to compute such an ordering [37, 32]. Our approach handles more timesteps by not considering the individual elements contained within clusters and dividing the sorting procedure into N individual blocks of T timesteps. To reduce the computational complexity—to achieve the least start-up-time of 40-60 seconds and to scale to up to 60 timesteps—our heuristic solution (barycenter approach [34])

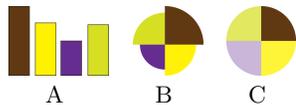


Figure 4: Design choices for visualizing cluster membership along with brain electrical activity. A

sweeps horizontally across all the clusters over N blocks of T timesteps (where NT is the total number of timesteps in the data) in a front-to-back and back-to-front manner to optimize the order and position of these communities on the vertical axis. This procedure results in a cluster ordering that minimizes the link distance between the timesteps.

4.1.3 *K*-Cluster Heatmap

Cluster analysis results can be sensitive to noise and prone to overfitting [21]. The *K*-Cluster heat map produces important information for the evaluation of the uniqueness of the number of clusters detected per timestep. Consensus clustering uses a cumulative distribution function (CDF) to determine an appropriate number of clusters K . In most cases, the likelihood for a single value of K will be large compared to the others, and the confidence that the chosen number of clusters is correct is high. The *K*-Cluster heat map (Fig. 2) shows the likelihood for a range of values of K for each time step. Black denotes high likelihood and white low likelihood. Using this heat map, analysts can identify timesteps where multiple values for K are almost equally likely and where confidence in the clustering results is low.

4.2 Electrode View

The electrode view shows cluster membership and electrical activity in a spatial context (G3) (Fig. 6 A,B), enabling the user to identify important spatial cluster evolution patterns (Q1, Q2 and Q3). These evolution patterns include 1) *spatial cluster distribution*, i.e., a cluster originally comprised of spatially adjacent electrodes splits into disjoint parts, 2) *spatial cluster combination*, i.e., a cluster consisting of disjoint regions becomes spatially coherent, and 3) *spatial activation*, i.e., the electrical activity of electrodes increases over time. The electrode view places interactive glyphs—representing electrodes—on a 2D subject-specific reconstructed brain model to provide the spatial context.

To support exploration at multiple temporal scales (G2) and effective comparison of cluster characteristics—i.e., cluster membership and electrical activity of the individual electrodes—over the spatial domain, our visualization method aggregates N multiples of continuous timesteps on each electrode view. Furthermore, interactive techniques that can be applied to glyphs, make it possible for the user to define the N multiples to be aggregated on each glyph.

The aggregation techniques provide the user with a detailed overview of the underlying data while not being overwhelmed by the entire data [9]. However, aggregated views have a drawback: they can lead to cognitive information overload. Nonetheless, for the multiples we consider, we found that our visualizations are better suited for the tasks performed by our collaborating neuroscientists.

4.2.1 Glyph Design



Figure 5: Varying the number of timesteps displayed on our glyph design.

All our glyph designs display n user-defined timesteps per individual electrode view, to save presentation space and to facilitate comparison between timesteps. We considered three visual designs for our glyphs, following some of the data aggregation guidelines specified by Elmqvist et al. [9]. The first design uses vertically stacked bar charts. Each bar represents one time step with its color indicating cluster membership and its height corresponding to electrical activity (Fig. 4A). The second design uses a clock metaphor [19] and subdivides each glyph into n equal slices. Each slice—starting at the top in clockwise order—represents a time step with its indicating cluster membership and its radius representing electrical activity (Fig. 4B). The third design is a variation of the second design and uses slice opacity instead of radius to represent the electrical activation level.

The first design depicts changes in cluster membership intuitively, but requires a large amount of glyph space. In the second design, the varying glyph sizes in the entire spatial layout impede the neuroscientists to assess the relative positions of the electrodes. Overall, our domain science collaborators preferred the third design where the glyph shape remains constant and only color and opacity are used to convey information. We use this design in our system and the remaining discussion in this paper is based on it.

4.2.2 Hierarchical Exploration

Patterns in brain activity occur at different temporal scales, where the appropriate scale may not be known *a priori*. To facilitate discovery of these patterns and scales, our tool controls the timesteps that can be displayed in a single electrode view. At the extremes, the method either displays all timesteps in a single electrode view (low granularity at the top level) or each time step in a separate electrode view (high granularity at the bottom level). Between these extremes, different levels of aggregation are possible. Different approaches exist to search for temporal evolution patterns. *Top-down* analysis starts with all timesteps in a single view and decreases aggregation until a pattern of interest is found. A *bottom-up* approach starts analysis with each time step in a separate view and increases aggregation until a pattern is found. However, exploration may also start at *mid-level*, in situations where the user already has a notion at what temporal scale a pattern may be identified.

5. CASE STUDIES

We demonstrate the usefulness of our tool by using both synthetic and real-world (seizure) datasets. The visual analysis approach greatly supports the comprehension of salient spatio-temporal evolution patterns and provides insight into the life-span of brain states and the clustering stability of the classification algorithm. A video demonstrating our tool in use can be found at <https://www.dropbox.com/s/icrm7pg1pr5r7hp/BrainKDD.mp4?dl=0>

5.1 Synthetic Dataset

We generated a simple synthetic dataset to better understand and evaluate our proposed methodology. In our data generation model, the electrodes had three known modes of normalized electrical activity, known intervals of activation patterns, and known likelihood values for the K-cluster heatmap. The controlled parameters of the data allowed us to investigate the features of the visualization outside the context of noisy brain recordings.

5.1.1 Spatio-temporal Analysis

We start our exploration with the evolution view, looking for general evolution trends in the data. In this view, three distinct clusters (colored in red, green and blue in Fig. 6) emerge and remain stable throughout timesteps [0-9] (in Fig. 6B). These clusters become randomly distributed at [10 to 19] to regain their stable configuration at [20 to 30].

To further examine the spatial configuration of these patterns, we employed a top-down approach exploring various temporal scales to identify a consistent activation pattern (similar activation patterns in one electrode view) across all electrode views. At a granularity of ten (ten-time points in one electrode view, Fig. 6A), persistent electrical patterns in each electrode view were found, e.g., glyphs in electrode views *one* and *three* were fully activated and deactivated, respectively. The electrode view *two*, on the other hand, showed a combination of activated and unactivated patterns.

To explore the intricate low-level activational and modular patterns that caused the temporal state change from unactivated to activated state, the granularity of the system was reduced to one (Fig. 6B). When examining timesteps [9, 10, 11], an emergent focal activation point (annotated in Fig. 6C) in the lower-right corner of the electrode view was evident. Further examination of subsequent timesteps revealed the progressive dominance of the red cluster over the region (Fig. 6C, views 9, 10 and 11). In summary, our combined visual analysis approach helped us categorize temporal states and identify low-level changes and their dependencies with high-level state changes.

5.2 Epileptic Seizure Dataset

Epilepsy is a neurological condition where the normal functioning of the brain is disrupted due to sudden bursts of electrical activity emanating from a certain region of the brain, i.e., seizure-initiating foci. This disruption is characterized by changes in the brain’s modular organization over time [24]. Exploring these differences may provide insight into the genesis and development of the seizures over time [24]. The neuroscientists on our team are primarily interested in: 1) identifying the brain regions that are active during various stages of the seizure and 2) identifying distinct spatio-temporal evolution patterns that characterize the onset and propagation of the seizure.

Data: The raw signal data from the ECoG electrode array was statistically analyzed to provide two spatio-temporal graph datasets using different pre-processing steps: a high-gamma dataset at a frequency ranging from 70 to 170 Hz, capturing multi-unit neuronal spiking, and a full-range dataset, averaging all frequencies captured by the recording device. We derived communities independently at each timestep using the consensus clustering algorithm [21].

5.2.1 Spatio-Temporal Analysis:

We now discuss some of the major insights obtained with our system by our collaborating neuroscientist and the neurosurgeon on our team through our tool:

Detection of brain states: Using the bottom-up approach, around a temporal scale of ten, neuroscientists found consistent progressive seizure activation patterns across all electrode views (Fig. 7C). Based on these patterns, the neuroscientists categorized *the electrode views into four major distinct brain states, i.e., before-seizure, early-seizure, mid-seizure, late-seizure.*

Detection of transitions between brain states: Based on the gradual changes in activity patterns (2200-2600ms in Fig. 7A), the seizure initiation zone (2200ms, Fig. 7A) (likely cause for state change from *before-seizure* to *early-seizure*) was found. The neuroscientists stated, *“the activated region is referred as Lateral Temporal Cortex (2400ms in Fig. 7A, and 2400ms in Fig. 8A). Through the glyph designs, activity patterns emerging from this point dynamically propagate to other parts of the brain (3000ms-5000ms in Fig. 7C).”*

Compare and contrast the evolution patterns governing different brain states: The signatures of cluster evolution at three distinct brain states (before-seizure, transition-state, late-seizure) were of interest to our users.

1) *800-1100ms:* Evolution patterns here appear to be stable and organized (highly confident clusters with a clear maxima in the K-cluster heatmap), with the same number of elements in each community (Fig. 8A, B). The spatial organization of clusters at 900ms (Fig. 8A) are scattered over the biological layout and are unactivated. The neuroscientist added, *“this distribution of clusters might indicate normal activity of the brain where specialized modules coordinate to perform a task.”*

2) *2400-2700ms:* Evolution patterns from this point onwards seem to become irregular. There is a noticeable activation of certain electrodes in spatial view and reduction in number of clusters. The neuroscientist commented, *“The transitions in the cluster evolution view seem to have a mix of irregular distribution of thin and thick links between timesteps. For example, the brown cluster in Fig. 8A and Fig. 8B seem to have comparatively less redistributions of electrode elements over time.”*

3) *5200-5500ms:* In the cluster evolution view, a dominant orange cluster (starting at 4000ms) emerges with few isolated small sized communities. The zero likelihood in the heatmap (interval 5500-6000ms) is caused as the algorithm produces only identity (1) values in the consensus matrix, showing no progressive change in the cumulative density plot (section 4.1.4), eventually picking the K value to be one.

Assess changes in community membership: In Fig. 7C, when evaluating the last electrode view, *there also seems to be a dominance of a single orange cluster over a significant period of time and space.*

When comparing the evolution patterns of the high-gamma and full-range pre-processed datasets (Fig. 8A and Fig. 8B), the domain experts revealed that, *“In general, the full-range dataset exhibits larger number of communities in all three sampled intervals (A,B,C). Although a single orange cluster is emergent in both datasets, it is only clearly visible in the high-gamma dataset. This pattern may be due to the high spatial locality and the high specificity of the frequency (70 to 170 Hz) used for the pre-processing of the data.”* The neuroscientist further added that *“The combination of time-*

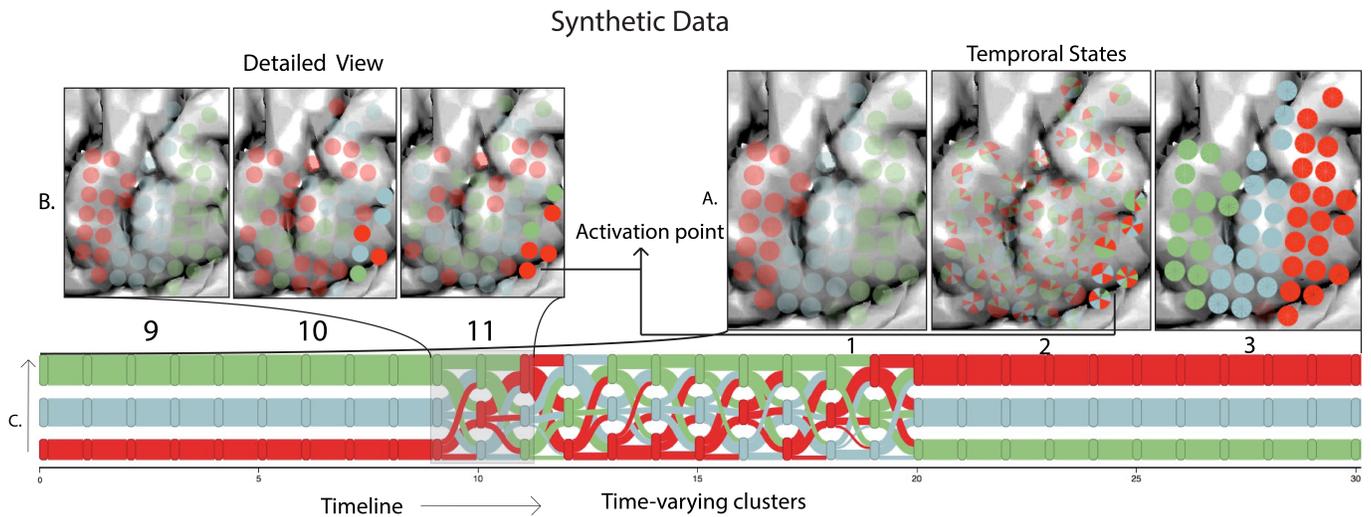


Figure 6: The figure shows evolution patterns underlying our generated dataset which we use to test and evaluate our approach. Color indicates the cluster configuration at each timestep and the opacities of the glyphs its electrical activity. A. Categorizing different temporal states, i.e., unactivated (view 1), transitional (view 2), activated (view 3). B. Evolution of the cluster assignment changes through the cluster evolution view, stable cluster configuration from timestep intervals [0-10] and [20-30], and, random unstable assignment over timestep interval [10-20]. C. Detailed analysis of a time interval selected in the cluster evolution view, activation patterns can be seen in the lower-right corner of the views for the timestep interval [10,11].

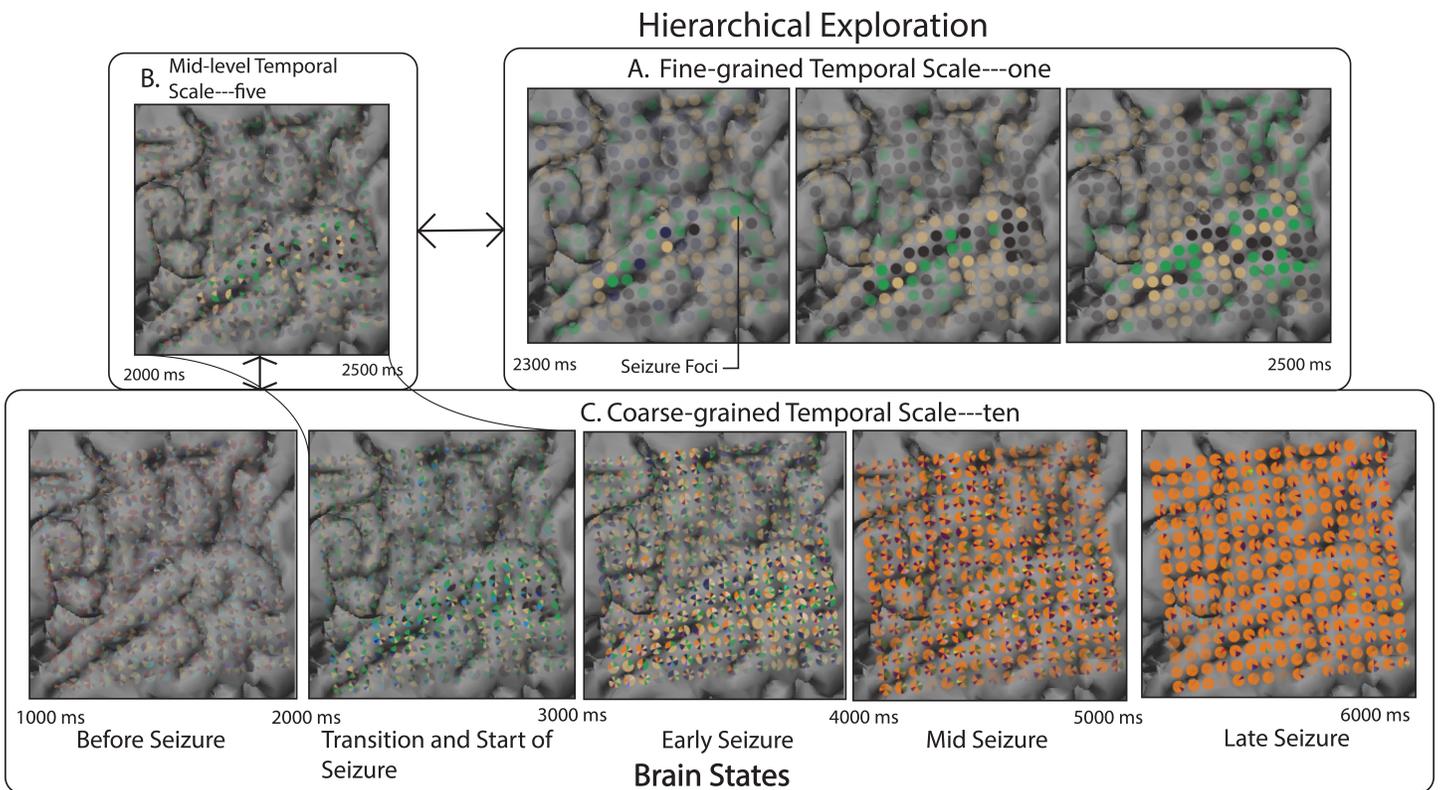


Figure 7: Hierarchical exploration of the seizure dataset at varying levels of granularity revealing the various brain states and associated cluster characteristics. Four major brain states are found at C. Upon further examination of their detailed evolution patterns in B and A, provides insight into seizure genesis and the initiation phase. The green cluster (in seizure initiating foci, i.e., ‘Lateral Temporal Cortex’) in A seems to play a prominent role in the seizure initiation phase (2200ms-2500ms).

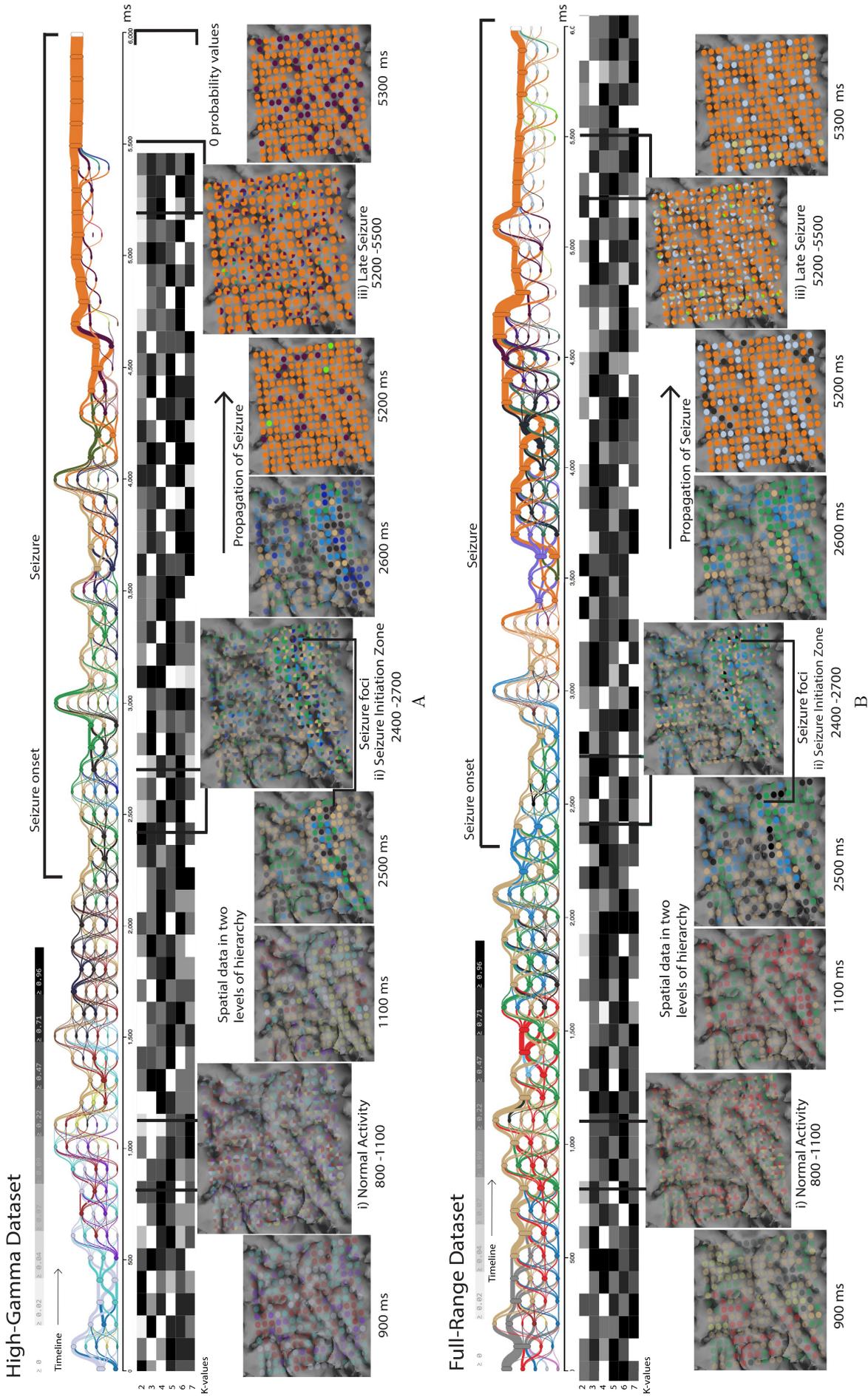


Figure 8: Visualizations of the two pre-processed datasets of the same raw input signals. Colors in the cluster evolution view and the electrode view indicate the time-varying communities. Three intervals are chosen for analysis to explore the modular signatures at different brain states. Significant evolution patterns at these stages are 1) normal activity—stable cluster characteristics with equally sized modules, 2) transition state—distinct activation patterns emerging from the ‘Lateral temporal cortex’, 3) late seizure state—single dominance of the orange cluster. When comparing the modular signatures between images A and B, the number of clusters in B seems to be larger

varying community structure characterization and the hierarchical multi-scale exploration approach has enabled these multi-part observations for the ECoG spatiotemporal data”.

In a short period, the neuroscientists could quickly identify various patterns of interest in the complex high-resolution dataset. These patterns are promising to reveal relevant insight into seizures, both their overall characterization and especially their genesis.

6. DISCUSSION

Neuroscientists analyzing and visualizing spatial-temporal graph data commonly use juxtaposed small-multiples or animations of node-link diagrams with a corresponding spatial view (brain maps). We have compared our approach with these baseline visualizations methods and point out how we gain new insights through our interactive system.

Small-multiples of brain maps: To identify low-level community membership changes of a node over time, information in each discrete timestep and its corresponding brain map must be examined. Such static multi-view approaches make it difficult to quickly assess the stability of nodes over a local spatial region of interest. Furthermore, to identify brain states with consistent activation patterns and investigate the global cluster lifetime phenomena, nodes have to be mentally grouped and compared based on their data attributes, further increasing the cognitive load for analysis.

Animations of brain maps: Animations are effective in capturing low-level spatial changes of communities over a small period—a change in community membership or activation is denoted by a sudden change in color. However, the frequently changing data attributes make it challenging to infer meaningful evolution patterns over long time-scales. This complication is due to its reliance on short-term memory that requires users to remember previous static views and then manually identify and compare the relevant changes. The same holds true for identifying brain states or global cluster evolution patterns.

Our approach: The identification of community membership changes of a node can be achieved by choosing an appropriate level of granularity and visually analyzing it over one static view (Fig. 7B and C). The analysis is reduced to finding the changes in the color of the consecutive slices in the glyph. Furthermore, the split-merge links in the cluster evolution view ensure that the nodes do not need to be re-identified at each timestep (small multiples) or remembered from previous timesteps (animations). Compared to these methods, interactive techniques in our approach help filter, select, zoom, pan and increase or decrease the level of granularity, helping gain a thorough understanding of the significant spatio-temporal patterns in the data. Unlike the small-multiples method, our method significantly reduces the visual scalability by aggregating temporal data over the spatial electrode view.

We now summarize our major findings resulting from exploring and analyzing ECoG data with our system.

Timeline-based representations enhance perception: The tightly integrated layout mechanism, where a consistent global timeline between the cluster evolution view and the electrode views is maintained, allows neuroscientists to correlate evolution patterns across the spatial layout and functional organization of the brain network. Furthermore, the side-by-side placement of the views on a timeline helps

the scientists compare and contrast salient patterns of interest for discontinuous timesteps.

Bottom-up analysis versus top-down analysis: The neuroscientists prefer different approaches for different tasks. The bottom-up analysis approach (Fig. 7, A-B-C) is chosen over the top-down approach (Fig. 7, C-B-A) for tasks relating to Q1 and Q2. As the bottom-up exploration method generates views with increasing complexity, relationships between evolution patterns at various scales can be seen clearly. In contrast, the top-down exploration method, which displays fewer electrode views with high levels of aggregation, can be cognitively overwhelming.

7. CONCLUSIONS AND FUTURE WORK

We have presented ECoG ClusterFlow, a hierarchical multi-scale approach for visualizing spatial and functional cluster evolution patterns. Our approach has allowed neuroscientists to investigate the major cluster evolution patterns over space and time. Through our approach, it is also possible to examine whether the major evolution events were a result of noise or a sudden change in functional or spatial properties of the network. Furthermore, we have discussed major neuroscience-driven data analysis tasks and design choices that led to the entire design of the system (done in close collaboration with neuroscientists) that helped gain insights for spatial cluster evolution data.

We plan to perform an evaluation concerning the usefulness of spatio-temporal clustering, i.e., a clustering algorithm that takes both the spatial as well as the temporal attributes into account. When utilizing both these attributes, interesting research questions arise, such as, “*is the resulting visualization an accurate representation of the data? How can we consistently compare such types of visualizations with a different dataset?*”. Moreover, we also aim to support scalable visual analysis methods for visualizing evolution patterns with more than 500 time steps

8. ACKNOWLEDGEMENTS

This work was supported through the Office of Advanced Scientific Computing Research grant “Towards Exascale: High Performance Visualization and Analytics Program,” program manager Dr. Lucy Nowell and by Laboratory Directed Research and Development (LDRD) funding from the Lawrence Berkeley National Laboratory, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

9. REFERENCES

- [1] J.-w. Ahn, C. Plaisant, and B. Shneiderman. A task taxonomy for network evolution analysis. *IEEE TVCG*, 20(3):365–376, 2014.
- [2] G. Andrienko, N. Andrienko, U. Demsar, D. Dransch, J. Dykes, S. I. Fabrikant, M. Jern, M.-J. Kraak, H. Schumann, and C. Tominski. Space, time and visual analytics. *Int. J. of Geographical Information Science*, 24(10):1577–1600, 2010.
- [3] G. Andrienko, N. Andrienko, M. Mladenov, M. Mock, and C. Politz. Identifying place histories from activity traces with an eye to parameter impact. *IEEE TVCG*, 18(5):675–688, 2012.
- [4] N. Andrienko, G. Andrienko, and P. Gatalsky. Exploratory spatio-temporal visualization: an

- analytical review. *J. of Visual Languages & Computing*, 14(6):503–541, 2003.
- [5] D. Archambault, H. C. Purchase, and B. Pinaud. Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE TVCG*, 17(4):539–552, 2011.
- [6] U. Braun, A. Schäfer, H. Walter, S. Erk, N. Romanczuk-Seiferth, L. Haddad, J. I. Schweiger, O. Grimm, A. Heinz, H. Tost, A. Meyer-Lindenberg, and D. S. Bassett. Dynamic reconfiguration of frontal brain networks during executive cognition in humans. *Proc. of the National Academy of Sciences*, 112(37):11678–11683, 2015.
- [7] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin. Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE TVCG*, 22(1):280–289, 2016.
- [8] P. Eades and M. L. Huang. Navigating clustered graphs using force-directed methods. *J. Graph Algorithms Appl.*, 4(3):157–181, 2000.
- [9] N. Elmqvist and J.-D. Fekete. Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE TVCG*, 16(3):439–454, 2010.
- [10] R. F. Erbacher, K. L. Walker, and D. A. Frincke. Intrusion and misuse detection in large-scale systems. *IEEE CG & A*, 22(1):38–47, 2002.
- [11] S. Fortunato. Community detection in graphs. *Physics reports*, 486(3):75–174, 2010.
- [12] P. Gatalaky, N. Andrienko, and G. Andrienko. Interactive analysis of event data using space-time cube. In *Proc. IEEE InfoVis*, pages 145–152, 2004.
- [13] D. Greene, D. Archambault, V. Belák, and P. Cunningham. Textluas: tracking and visualizing document and term clusters in dynamic text data. *Tech Report, arXiv:1502.04609*, 2014.
- [14] D. Greene, D. Doyle, and P. Cunningham. Tracking the evolution of communities in dynamic social networks. In *Proc. IEEE ASONAM*, page 176, 2010.
- [15] M. Harrower and C. A. Brewer. Colorbrewer.org: an online tool for selecting colour schemes for maps. *The Cartographic J.*, 40(1):27–37, 2003.
- [16] Y. He, J. Wang, L. Wang, Z. J. Chen, C. Yan, H. Yang, H. Tang, C. Zhu, Q. Gong, Y. Zang, and A. Evans. Uncovering intrinsic modular organization of spontaneous brain activity in humans. *PloS one*, 4(4):e5226, 2009.
- [17] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80:360–378, 2013.
- [18] M. Jern and J. Franzen. Geoanalytics-exploring spatio-temporal and multivariate data. In *Proc. InfoVis*, pages 25–31, 2006.
- [19] B. Kondo, H. Mehta, and C. Collins. Glidgents: Interactive glyphs for exploring dynamic graphs. In *Proc. IEEE InfoVis (Poster)*, 2014.
- [20] C. Meisel and C. Kuehn. Scaling effects and spatio-temporal multilevel dynamics in epileptic seizures. *PLoS One*, 7(2):e30371, 2012.
- [21] S. Monti, P. Tamayo, J. Mesirov, and T. Golub. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine learning*, 52(1-2):91, 2003.
- [22] S. Murugesan, K. Bouchard, J. A. Brown, B. Hamann, W. W. Seeley, A. Trujillo, and G. H. Weber. Brain modulyzer: Interactive visual analysis of functional brain connectivity. *IEEE TCBB*, PP, In press (2017).
- [23] C. Plug, J. C. Xia, and C. Caulfield. Spatial and temporal visualisation techniques for crash analysis. *Accident Analysis & Prevention*, 43(6):1937, 2011.
- [24] S. Ponten, F. Bartolomei, and C. Stam. Small-world networks and epilepsy: graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures. *Clinical neurophysiology*, 118(4):918–927, 2007.
- [25] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, and S. Petersen. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.
- [26] H. C. Purchase, E. Hoggan, and C. Görg. How important is the mental map?—an empirical investigation of a dynamic graph layout algorithm. In *Int. Sym. on GD*, pages 184–195, 2006.
- [27] K. Reda, C. Tantipathananandh, T. Berger-Wolf, J. Leigh, and A. Johnson. Socioscape—a tool for interactive exploration of spatio-temporal group dynamics in social networks. In *In Proc. IEEE InfoVis*, 2009.
- [28] K. Reda, C. Tantipathananandh, A. Johnson, J. Leigh, and T. Berger-Wolf. Visualizing the evolution of community structures in dynamic social networks. *CGF*, 30(3):1061–1070, 2011.
- [29] F. Reitz, M. Pohl, and S. Diehl. Focused animation of dynamic compound graphs. In *Proc. IEEE InfoVis*, volume 13, pages 679–684, 2009.
- [30] P. Riehmann, M. Hanfler, and B. Froehlich. Interactive sankey diagrams. In *Proc. IEEE InfoVis*, pages 233–240, 2005.
- [31] M. Rosvall and C. T. Bergstrom. Mapping change in large networks. *PloS One*, 5(1):e8694, 2010.
- [32] A. Sallaberry, C. Muelder, and K.-L. Ma. Clustering, visualizing, and navigating for large dynamic graphs. In *Int. Sym. on GD*, pages 487–498, 2012.
- [33] O. Sporns and R. F. Betzel. Modular brain networks. *Annual review of psychology*, 67:613–640, 2016.
- [34] K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical system structures. *IEEE Trans. on Systems Man and Cybernetics*, 11(2):109–125, 1981.
- [35] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko. Stacking-based visualization of trajectory attribute data. *IEEE TVCG*, 18(12), 2012.
- [36] E. R. Tufte. Envisioning information. *Optometry & Vision Science*, 68(4):322–324, 1991.
- [37] C. Vehlow, F. Beck, P. Auwärter, and D. Weiskopf. Visualizing the evolution of communities in dynamic graphs. *CGF*, 34:277–288, 2015.