Visualizing Routine Dynamics in Outpatient Medical Clinics with Topological Data Analysis*

Li Zhang[1], Julie Ryan Wolf[2], Alice P. Pentland[2], and Brian T. Pentland[1][0000−0001−7526−2476]

[1] Eli Broad College of Business, Michigan State University, East Lansing MI 48824, USA
zhan2070@msu.edu, pentland@broad.msu.edu

[2] School of Medicine and Dentistry, University of Rochester, Rochester NY 14642, USA
julie_ryan@urmc.rochester.edu, Alice_Pentland@urmc.rochester.edu

Abstract. We demonstrate the use of topological data analysis (TDA) to visualize the dynamics of organizational routines in outpatient medical clinics over the course of the COVID-19 pandemic. We use Electronic Health Record audit trail data from January 2020 to December 2021 to visualize changes in routines in outpatient clinics in four different medical specialties. By representing the pattern of action each day in each clinic as a weighted, directed graph, we see that some clinics bounced between several distinct patterns of action over the two year period. In contrast, other clinics never bounced back to the pre-COVID pattern of action. In all of the clinics, we see evidence of temporal auto-correlation: the pattern of action on any given day is similar to the pattern of action on days that are close in time. Because it can capture recurrence of action patterns that might be widely separated in time, TDA offers a substantial advance in the state of the art for visualizing the dynamics of organizational routines.

Keywords: Drift · Change · COVID-19 Pandemic · Electronic Health Records · · Routine dynamics.

1 Introduction

Routines can change on their own, but sometimes they are pushed. For example, the COVID-19 pandemic provided a strong external shock that disrupted normal processes and routines, especially in healthcare organizations [4]. Following

* This research was supported by the National Science Foundation under Grant No. BCS-2120530. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This research was also supported in part by University of Rochester CTSA (UL1 TR002001) from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health (NIH). The content is solely the responsibility of the author(s) and does not necessarily represent the official views of the National Institutes of Health.

guidance from the Centers for Medicare and Medicaid Services (CMS)[3] issued on March 18th, 2020, many clinics suspended their office visits entirely. Of course, many patients still had medical conditions that required attention, so healthcare providers scrambled to devise alternative ways to treat their patients. Standard procedures and routines needed to be reinvented to protect the patients and the clinical staff.

In this article, we use the example of the COVID-19 pandemic to demonstrate the use of topological data analysis (TDA) for visualizing recurrent patterns of action in organizational routines. We examine changes in routines in four outpatient clinics connected to a large U.S. medical center from January 2020 through December 2021. We build on the approach introduced by Pentland et al. [6] to visualize change in patterns of action over time. This approach is based on the idea of representing patterns of action as weighted, directed graphs and comparing the graphs over time. Here, we demonstrate the additional insights gained from topological data analysis (TDA) and discuss their implications for the theory of routine dynamics. Our findings show evidence of temporal autocorrelation in all the clinics: clinic-days that are close in time tend to have similar patterns of action. While changes from day to day are incremental, they are also systematic: the clinics move between distinctly different, recurrent patterns of action.

TDA describes a broad category of methods motivated by the intuition that the *shape* of data contains useful insights about the underlying phenomenon [1]. One particularly significant shape is the cycle (or circle or loop), which indicates recurrence (return to a prior state). Recurrence is fundamental to organizational routines [3], organizational path dependence [8], and institutions [5]. Here, we adapt a tool called Temporal Mapper [11] that uses TDA to visualize and analyze high-dimensional networks that change over time. Originally developed for the analysis of neurological networks, Temporal Mapper is a general purpose tool for identifying recurrent structures in temporal networks.[4] This perfectly describes the problem of analyzing routine dynamics using digital trace data.

2 Electronic Health Record Audit Trail Data

Audit trail data were extracted from the EPIC Electronic Health Record (EHR) system. Audit trail data is meta-data about the record keeping process, so it does not contain any clinical data. It provides an accurate view of who accessed or updated the clinic records, but it provides a limited view of the actual medical work.

The data we report here comes from four clinical areas: cardiology, dermatology, oncology and radiation oncology. The data cover a two year window from January 1, 2020 through December 31, 2021. This two year period spans the most severe part of the COVID-19 pandemic, which hit the Northeastern United States in March 2020.

[3] https://www.cms.gov/files/document/covid-elective-surgery-recommendations.pdf
[4] Code is available here: https://github.com/braindynamicslab/tmapper.

3 Visualizing Routine Dynamics with TDA

Audit trail data provides a rich source of information about patterns of action in the EHR record keeping process. The method we demonstrate here builds on the approach introduced by Pentland et al. [6] to visualize change over time by using a time-series of weighted, directed graphs (also known as directly follows graphs, or DFGs). This network time series is used in TDA to map the trajectory of clinical routines over time.

3.1 A time series of weighted, directed graphs

We describe the pattern of action each day in each clinic as a weighted, directed graph. This graph provides a description of the "state" of the clinic on any given day and provides a way to compare any two clinic-days. The vertices in the graphs are the actions recorded in the audit trail, the edges are the sequential relations between actions within each patient encounter, and the weights denote the frequency of the sequential relations between actions.

We compute the graph for each day in each clinic from 7am to 7pm. This is a reasonable choice because the clinics only operate during the day and shut down at night. We compute the graph without filtering the data. This maximizes fitness without regard for any other metric of model quality. The resulting networks are very complex, but they provide a sensitive indicator of change.

As described by Pentland et al.[6], we use cosine distance based on the frequency of the edges in the graphs to compute the distance between clinic-days. The distance (or similarity) between clinic-days is an important input for topological data analysis (TDA). We use the graph for each clinic-day to represent the state of the clinic on that day and we analyze the trajectory of each clinic over time. This approach departs from methods where drift is detected or measured relative to a known process model. This approach is model-free; it is based on recurrence in the unfiltered DFG.

3.2 Simple time series

Pentland et al.[6] suggest two ways to plot the data based on the time-series of networks for each clinic-day. One way shows the change from prior day. This plot answers the question, "Is today the same as yesterday?" The second way shows the change from the pre-pandemic pattern of action. This plot answers the question, "How does today compare to a baseline?"

Figure 1 illustrates both of these visualization for one of the dermatology clinics. We smooth the plots and include a 95% confidence interval highlighted in gray. The plot on the left of the figure shows a comparison to the previous day, and the plot on the right of the figure shows a comparison to the pre-pandemic baseline. The pattern of action changes a lot and does not appear to bounce back. If anything, it appears to have stabilized into a "new normal" that is different from the original, pre-pandemic pattern of action.

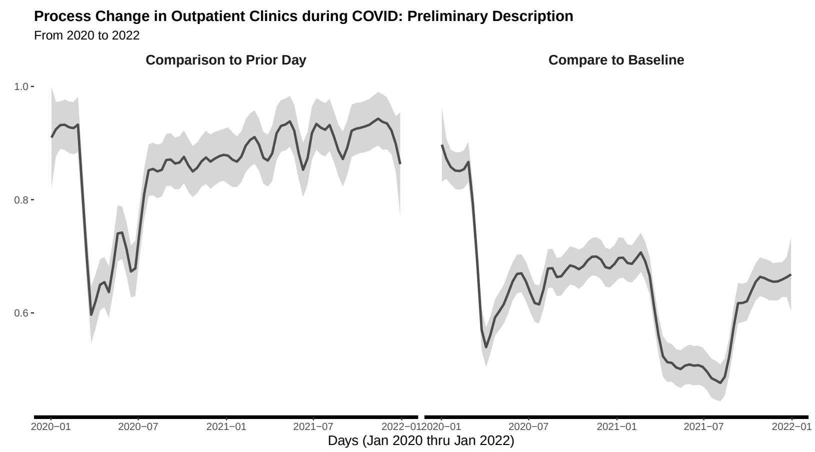


Fig. 1: Change over time in one dermatology clinic

3.3 Classic recurrence plot

Figure 1 is a helpful visualization of process change, but we can extract and visualize lot more information from the same data using a classic recurrence plot, as shown in Figure 2. It is called a recurrence plot because it indicates recurrent states (similar patterns of action) over time. The plot in Figure 2 shows the pair-wise distance between *every* pair of clinic-days for the entire two-year time-series for one dermatology clinic. By comparison, the right-hand side of Figure 1 shows only one row from Figure 2 (comparison to a baseline). The left-hand side of Figure 1 shows only the diagonal (offset by one, so it shows comparison to the previous day). Figure 2 is symmetric because the cosine distance measure is not directional.

The shading in this recurrence plot provides an intuitive picture of how the pattern of action changed over time. Dark areas show similar patterns of action; light areas show different patterns of action. For example, the small dark square in the upper-left corner is the pre-COVID routine, which was suddenly disrupted in March 2020. There is a significant period of moderate change, followed by two more rather sudden, substantial changes. The timing of these changes corresponds to the changes on the right-hand side of Figure 1. Throughout the recurrence plot, there are bright lines that indicate days with exceptional patterns of action.

3.4 Recurrence and sequence of states using TDA

TDA allows us to visualize the same data in a way that shows the temporal progression of the pattern of action. It allows us to see patterns of recurrence that are not evident in Figures 1 or 2. In effect, it tells a visual story of how the patterns of action are related in terms of similarity and time.

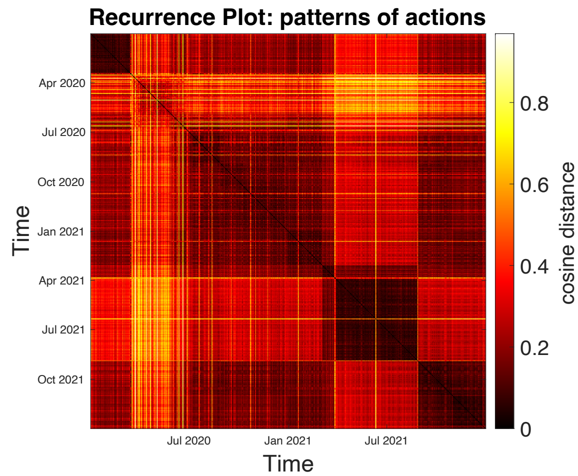


Fig. 2: Classic recurrence plot showing change over time in dermatology

3.5 Temporal Mapper Algorithm

To visualize the trajectory of the pattern of action, we use a TDA tool called Temporal Mapper [11]. The Temporal Mapper algorithm has four stages.

1. Cluster similar states using kNN. In the first stage, Temporal Mapper uses a k-nearest neighbor (kNN) algorithm to cluster clinic-days based on similarity (as measured by cosine distance). In principle, other clustering methods are possible, but KNN is ideally suited to the problem of identifying recurrence. For kNN clustering, the algorithm uses Figure 2 as input. For the plots we show in this article, we set $k = 5$. This choice of k assumes that, on average, there are five non-adjacent work days that share the similar pattern of actions during the two year period. This assumption balances the inherent variability of organizational routines against their tendency to exhibit stability [2]. Lower values of k yield finer grained results (more clusters); higher values of k yield coarser grained results (fewer clusters).

The kNN algorithm results in a network of states with several clusters or network components. Each cluster contains a set of similar states. In this network, nodes represent patterns of action, and the edges represent similarity (as in Figure 2). In this network, the edges are un-directed (or bi-directional). Even if clinic-days are widely separated in time, they can be clustered together as the same state if they are sufficiently similar.

2. Neighborhood graph. In the second stage, the algorithm introduces the temporal sequence back into the picture, forming a network with two types of edges: temporal edges and similarity edges. Temporal edges are directed and similarity edges are un-directed. In Temporal Mapper, this is called the neighborhood graph. Nodes can be neighbors in time, in similarity, or both.

3. State transition graph. In the third stage, the algorithm merges any circles formed by the directed edges that are shorter than a preset threshold. We set the threshold $d = 2$. This means that, for any two nodes a and b, we merge them together if the path lengths from a to b and from b to a are both less than or equal to 2. As the edges in stage 1 are all reciprocal, the distance between any two similarity neighbors is 1; therefore, the nodes in each cluster merge as a larger node. In this stage, the algorithm just continues to merge all circles under the threshold. After merging, the only visible edges are temporal edges: the "arrows of time". The resulting graph is called a state transition graph, showing the trajectory of the pattern of action as it changes over time. In this network, each node represents a "state", and each edge represents an "arrow of time."

4. Geodesic recurrence plot. Finally, using the state transition graphs, the algorithm calculates the pair-wise path lengths between each two clinic-days. In the state transition graphs, one node can contain multiple clinic-days, and the path lengths between two states are the shortest distances between clinic-days in the two states. The classic recurrence plot in Figure 2 is based only on similarity (measured by cosine distance), but the geodesic recurrence plot is based on the distance between states in the state transition graph, which includes information about time and similarity.

3.6 Results of Temporal Mapper Visualization

Temporal Mapper produces plots that show the trajectory (or temporal progression) of the routine from one state to the next [7]. In this section, we show one example plot from each clinical specialty. In these plots, the size of the nodes indicates the number of times that state is observed in the data. In addition to the state transition graph, the geodesic recurrence plot shows distance between states on the path (or trajectory). Darker color indicates that the states are close together; light color indicates that the states are farther apart.

In all four clinics, the pre-pandemic "baseline" condition in January and February 2020 shows up as a small, dark box in the upper-left corner. When the pandemic hit in March 2020, we can see varying levels of reaction in the four clinics. Intuitively, one might expect more dramatic differences. In dermatology, office visits were reduced by 96% during the initial lockdown. Other clinical areas continued to see patients, but at a reduced rate. However, many differences in clinical practice (e.g., social distancing, use of masks and gloves, etc.) do not show up in the electronic health record because they did not change the clinical documentation process. In contrast, in all four clinics, changes in guidance from the Centers for Medicare and Medicaid Services (relating to the availability of COVID-19 vaccines, starting in March 2021) is readily visible. While each medicial specialty implemented this guidance in relation to their own caseload, it affected the documentation process in all four clinics. For example, clinical staff had to visit different screens and enter different information relevant to the vaccination status of the patient. As with any kind of digital trace data, we can only see what the EHR sees.

Cardiology (Figure 3) In Cardiology, the trajectory has two main states. We can see that the pattern of action was affected by the initial lockdown in March 2020, but it bounced back fairly quickly. A much larger change occured later, in March 2021, when COVIC vaccines became available for different groups of patients. Eventually, the routines in the cardiology clinic have "bounced back" because the pattern of action is similar to the original, pre-pandemic pattern of action.

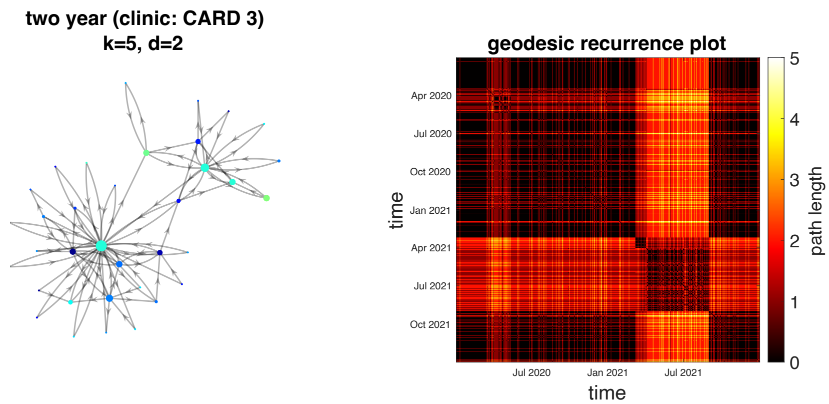


Fig. 3: Trajectory of Routine Cardiology Clinic

Dermatology (Figure 4) In Dermatology, we can see a clear trajectory of distinct states. According to our co-authors (AP Pentland and J Ryan Wolf), Dermatology increased the use of telemedicine and implemented a novel "eConsult" service, so that primary care physicians could get quick consultations for patients with skin problems. As a result of these changes, the pattern of action continued to get more distant from the original, pre-pandemic pattern. The white (blank) area in the geodesic recurrence plot means that the there is no path back to the prior state (no recurrence). COVID appears to have provided the impetus for a permanent change to the routines in this Dermatology clinic.

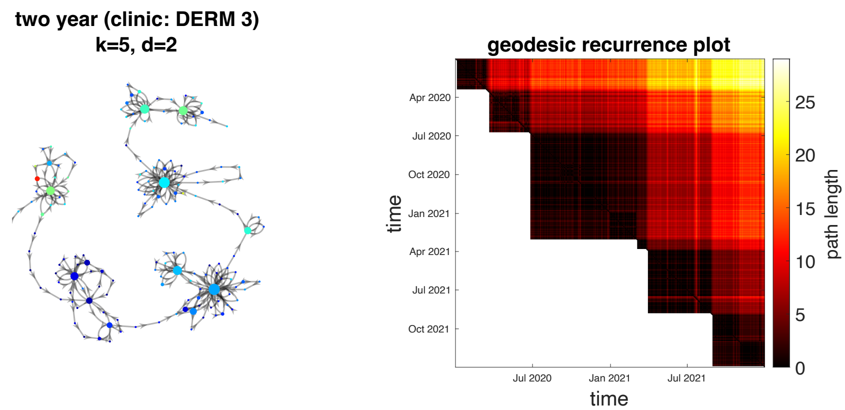


Fig. 4: Trajectory of Routine in Dermatology Clinic

Oncology (Figure 5) Like Dermatology, the trajectory in Oncology has several distinct states. Unlike Dermatology, the trajectory bounces around between states and eventually returns to a state that is close to the original, pre-pandemic pattern of action. As in all of the other clinics, we can see a distinct change in the pattern of action when COVID-19 vaccines became available in March 2021.

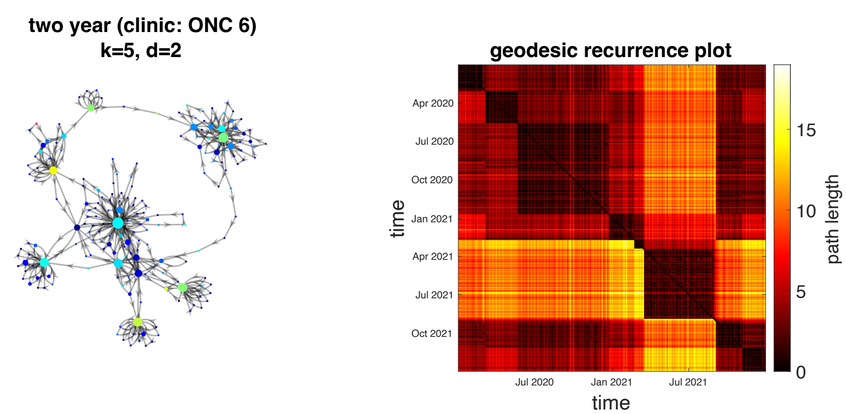


Fig. 5: Trajectory of Routine Oncology Clinic

Radiation-oncology (Figure 6) In this clinic, there are three distinct groups of states. Unlike the other clinics, the trajectory is not very much affected by the lockdown in March 2020. This may be because the patients in Radiation Oncology were receiving treatments that had to be provided on a strict schedule. Over the 2 year period, the clinic moved through three different states. By the end of 2021, the pattern of action in Radiation Oncology was similar to the pre-pandemic pattern, but not identical.

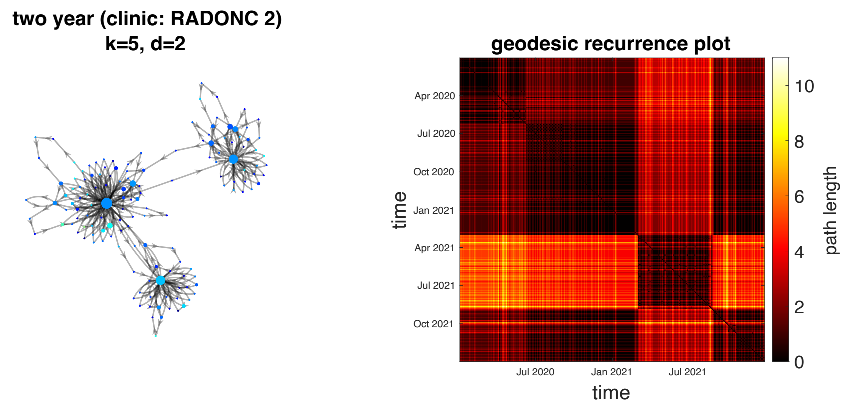


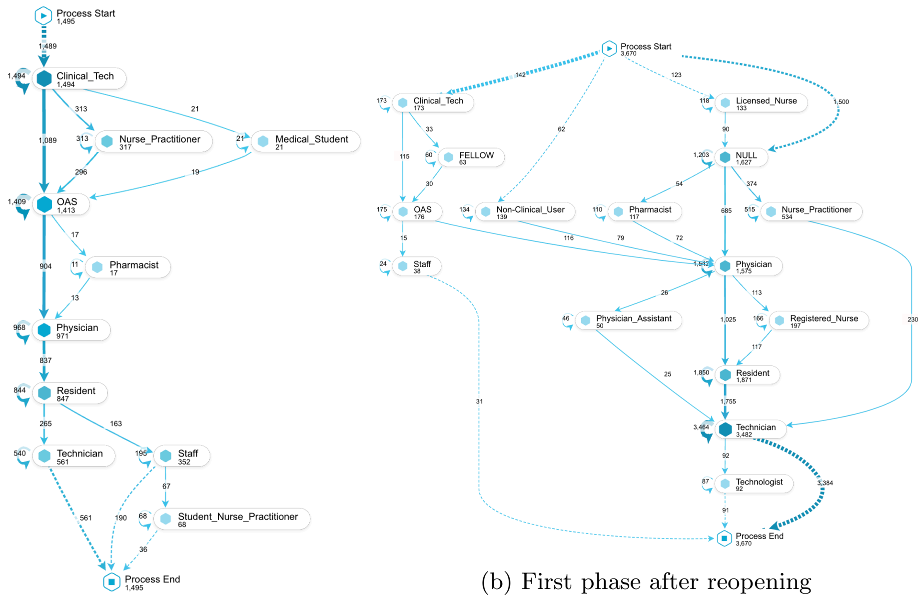
Fig. 6: Trajectory of Routine in Radiation Oncology Clinic

3.7 Visualizing the pattern of action in different states

Visualizing the trajectory can help facilitate explanations, as well. For example, we can drill down into each of the states in the trajectory of the routine by visualizing the pattern of action when the clinic is in that state. To illustrate differences between states, we used Celonis to generate process maps for the pattern of action in two distinct states in the Dermatology clinic (see Figure 7). We filtered the data to create plots that are easier to interpret. The first state is during the time period from January to March, 2020 (i.e., pre-pandemic). The second state is during the time period from July 2020 to February, 2021 (i.e., first phase after reopening). From Figure 7, we can see that the processes in the two states are quite different. For example, before pandemic, Clinical_Tech performed 100% of the check-in work; however, during the first phase of reopening, only 4% of the check-in work is performed by Clinical_Tech.

4 Discussion

Feldman et al.[3] stated that "Routine Dynamics focuses on tracing actions and associations between actions" (p. 506). Focusing on actions within a routine has opened up many avenues of inquiry and insights about how routines work. Here, we are showing how Routine Dynamics can expand its focus to trace the trajectory of a routine as it changes over time. TDA provides a way of zooming out, so we can trace *patterns* and associations between *patterns*.



(a) Pre-pandemic

(b) First phase after reopening

Fig. 7: Comparing two states in the Dermatology clinic

Of course, tracing patterns with digital trace data has limitations. In clinical settings, EHR data provides a shadow of the actual medical work. Neurological research has a similar limitation. We can detect electrical signals in the brain, but it's impossible to know what a subject is really thinking. The audit trail data we analyzed does not contain any information about social distancing, personal protective equipment, or other changes in clinical practice. The audit trail data only shows actions in the health record, not the actual medical work.

The uncertainty (or noise) associated with trace data is evident in the visualizations we have presented here. In every clinic, there appear to be groups of closely related states – patterns of action that are similar but not similar enough to be grouped together. This may be an artifact of our choice to use "unfiltered" audit trail data to describe the patterns of action in each clinic-day. The data are noisy, and there is a lot of natural variability in clinical routines, so it is no surprise that the clinics appear to "bounce around" from day to day. At the same time, these examples illustrate the power of the Temporal Mapper algorithm to extract and display a clear temporal progression of action patterns.

4.1 Telling the story of change with TDA

This clarity of structure allows us to visualize the trajectory of change in a way that has never been possible before. We can detect drift, but we cannot understand its significance until we place in the context of a larger trajectory.

The significance of any particular change is not evident until we compare it to other changes. By taking into account the pair-wise similarity of all clinic-days (as in Figure 2), TDA uses all of the available information to construct a trajectory. Based on these trajectories, we see three important generalizations for routine dynamics.

Temporal auto-correlation in every trajectory. While the trajectory of each clinic has unique features, they all illustrate some amount of temporal auto-correlation: States that are adjacent in time tend to be similar. In geography, Tobler's First Law of Geography [9] is based on spatial auto-correlation. It states that spatially adjacent locations tend to have similar properties. Of course, this "law" has many exceptions, such as shorelines. Here, we have visual evidence for a First Law of Routine Dynamics: temporally adjacent patterns of action tend to be similar. Temporal auto-correlation implies continuity in patterns of action over time.

More than just incremental drift. While there is evidence of continuity, every clinic also shows clear evidence that the pattern of action jumps from state to state in discrete, identifiable transitions. It does not look like incremental drift or a gradual accumulation of minor, situation-specific improvisations, workarounds, exceptions and so on. There are day-to-day variations in all of the clinics, but there are also discrete transitions as the pattern of action changes from from state to state.

More than just exogenous shock. Some of the state changes we can see are clearly driven by external factors that affected all of the clinics at once. For example, guidance from the Centers for Medicare and Medicaid Services (in March 2021) appears to have a bigger effect than the onset of the virus itself (in March 2020). However, in spite of the fact that all of these clinics operated in the same hospital system, in the same regulatory context, and the same EHR software, the trajectory of every clinic is different. The Cardiology clinic has two distinct groups of states, but Radiation Oncology has three, and the other clinics have more. Three of the clinics bounce around between states, but Dermatology just keeps changing to new states. We clearly see the change introduced by vaccine availability in every clinic, but other aspects of the trajectory appear to be unique to each clinic.

4.2 Beyond Stability and Change: Visualizing Recurrence

Tsoukas and Chia [10] note that theorists tends to assume stability and then try to explain change. However, the patterns of action in medical clinics are highly variable. Every moment, in every clinic, patterns of action form and dissolve. There is continuous change. At the same time, as we can clearly see, the patterns of action are highly recurrent. There is continuous change, but also periods of stability.

A major theoretical advantage of the method presented here is that it does not assume stability or change; the only assumption is temporal progression, which is logically antecedent to both stability and change. Given a temporal progression of states, as suggested by Pentland et al [6], Temporal Mapper helps us identify and visualize recurrence, which is an indicator of both stability and change.

5 Conclusion

Topological Data Analysis (TDA) provides a new way to visualize and analyze recurrence in high-dimensional, temporal data. This kind of data is increasingly available for organizational routines. Because it provides a way to visualize recurrence, TDA can provide new avenues for inquiry and insight into the dynamics of organizational routines and other processual phenomena in organizations.

References

1. Epstein, C., Carlsson, G., Edelsbrunner, H.: Topological data analysis. Inverse Problems **27**(12), 120201 (2011)
2. Feldman, M.S., Pentland, B.T.: Reconceptualizing organizational routines as a source of flexibility and change. Administrative Science Quarterly **48**(1), 94–118 (2003)
3. Feldman, M.S., Pentland, B.T., D'Adderio, L., Lazaric, N.: Beyond routines as things. Organization Science **27**(3), 505–513 (2016)
4. Herrera, C.A., Juárez-Ramírez, C., Reyes-Morales, H., Bedregal, P., Reartes-Peñafiel, D.L., Díaz-Portillo, S.P., Klazinga, N., Kringos, D.S., Veillard, J.: Covid-19 disruption to routine health care services: How 8 latin american and caribbean countries responded: The study examines disruption to routine health care services brought about by the covid-19 pandemic in eight latin american and caribbean countries. Health Affairs **42**(12), 1667–1674 (2023)
5. Ocasio, W.: Institutions and their social construction: A cross-level perspective. Organization Theory **4**(3), 26317877231194368 (2023)
6. Pentland, B., Vaast, E., Wolf, J.R.: Theorizing process dynamics with directed graphs: A diachronic analysis of digital trace data. MIS Quarterly **45**(2) (2021)
7. Pentland, B.T., Yoo, Y., Recker, J., Kim, I.: From lock-in to transformation: A path-centric theory of emerging technology and organizing. Organization Science **33**(1), 194–211 (2022)
8. Sydow, J., Schreyögg, G., Koch, J.: On the theory of organizational path dependence: Clarifications, replies to objections, and extensions. Academy of Management Review **45**(4), 717–734 (2020)
9. Tobler, W.: A computer movie simulating urban growth in the detroit region. Economic Geography **46**(sup1), 234–240 (1970)
10. Tsoukas, H., Chia, R.: On organizational becoming: Rethinking organizational change. Organization Science **13**(5), 567–582 (2002)
11. Zhang, M., Chowdhury, S., Saggar, M.: Temporal mapper: Transition networks in simulated and real neural dynamics. Network Neuroscience **7**(2), 431–460 (2023)