Segmentations with Explanations for Outage Analysis

Liangzhe Chen*,°, Nikhil Muralidhar*,°, Supriya Chinthavali+,
Naren Ramakrishnan*, B. Aditya Prakash*
*Department of Computer Science, Virginia Tech
+Oak Ridge National Laboratory
Email: {liangzhe, nik90, naren, badityap}@cs.vt.edu, chinthavalis@ornl.gov
°(equal contribution)

ABSTRACT

Recent hurricane events have caused unprecedented amounts of damage and severely threatened our public safety and economy. The most observable (and severe) impact of these hurricanes is the loss of electric power in many regions, which causes the breakdown of many public services. Understanding the power outages and how they evolve during a hurricane provide insights on how to reduce outages in the future, and how to improve the robustness of the underlying critical infrastructure systems.

In this paper, we propose a novel segmentation with explanations framework to help experts understand such datasets. Our method, CUT-n-REVEAL, first finds a segmentation of the outage sequences to capture pattern changes in the sequences. We then propose a novel explanation optimization problem to find an intuitive explanation of the segmentation, that highlights the *culprit* of the change. Via extensive experiments, we show that our method performs consistently in multiple datasets with ground truth. We further study real county-level power outage data from several recent hurricanes (Matthew, Harvey, Irma) and show that CUT-n-REVEAL recovers important, nontrivial and actionable patterns for domain experts.

1 INTRODUCTION

Power outages during several recent hurricanes have caused severe impact on our national security, economy and public safety. The 2017 hurricane season was the most expensive in U.S. history resulting in huge economic losses (greater than $250 billion). Hurricane Irma caused one of the largest power outages which reportedly knocked out power to 4.5 million of Florida Power & Lightfis 4.9 million customers. Hence, better understanding the power outages and how they evolve during the hurricanes is a very important task for damage prevention and control.

Domain experts in critical infrastructure systems (CIS) constantly seek solutions and ideas on how to reduce the power outages during hurricanes. For example, Oak Ridge National Laboratory's (ORNL) Energy Awareness and Resiliency Standardized Services (EARSS) project developed a fully automated procedure to take wind speed and location estimates provided by hurricane forecasters and provide a geospatial estimate on the impact to the electric grid in terms of outage areas and projected duration of outages [2]. There are many such examples, including the National Infrastructure Simulation and Analysis Center (NISAC) program[1] (which provides projected outages) and ANL's HEADOUT model which quickly estimates potential customers who will lose power.

Although crucial, it is often non-trivial to analyze the cause of power outages. While the majority of power failures from national grids last only a few hours, some blackouts can last days and can impact several critical infrastructure systems like telecommunication networks, financial services, water supplies and hospitals. For e.g., from airlines to blood supply levels to energy and water supplies, the 2003 NE American blackout impacted a wide range of critical infrastructure and emergency management sectors in both Canada and the U.S. One of the widespread impacts of Hurricane Sandy was loss of electricity for over 8.5 million homes across the eastern US [3]. In Coastal Communities, loss of electricity led to the loss of refrigeration (and therefore caused food spoilage) and observed cascading consequences including the interruption of communication systems, disruption of heating and cooling systems, critical medical support equipment, water supply and sanitation facilities, other building mechanical functions, and government services.

As a result, identifying how and why the power outages evolve in a certain way during historic hurricanes is a very challenging task. Identifying time-points where there is a sudden change in the evolution of number of outages can help in many aspects from identifying causes, prioritizing repair resources, to predicting future outage outbreaks [7, 12]. Such a problem may be addressed using the time-series mining task of 'segmentation' (however there are some domain-based properties of such time-series which we need to take into account for segmentations). Nevertheless, merely providing cut-points via segmentation is usually not enough for directly actionable guidance. Knowing the time-point of changes, domain experts also need to know *what* is changing and *how* it is changing, which is typically deeply buried in an abundance of unrelated information. In the context of power outage data from hundreds of counties, helping power engineers understand which counties were responsible for a particular sudden change is very valuable. Knowledge of the so-called 'culprits' allows them to direct their attention to these counties, and perform faster and more targeted damage control, to prevent even larger failures.

In this paper we address this issue via a novel segmentation-with-explanations approach. Our main contributions are:

- We propose a segmentation algorithm for modeling power outages, that captures temporal relationships among different time stamps.
- We propose an explanation algorithm that automatically identifies the culprit of changes for the segmentation detected.
- Via experiments on historical hurricane data, we identify when and how the outage numbers dramatically change,

[1] http://www.sandia.gov/nisac/

and we summarize how these can help us reduce the number of outages in future disaster events.

The rest of the paper is organized as follows. We first provide a problem definition, followed by our solution and algorithm and finally our experimental results.

2 FOCUS AND SET-UP

We first briefly describe our focus and set-up. Large power grids usually contain thousands of generators, hundreds of thousands of transmission lines and millions of consumers. Grid components have strong interdependencies like in the transmission grid where multiple paths exist between generators and consumers and these paths typically are arranged in a mesh grid manner. Hence if one path or line fails, the electricity instantaneously follows an alternate path governed by Kirchhoff's voltage and current laws [11]. If the alternate path however cannot handle the overload in flow, it in-turn fails and this failure cascades to neighboring components. Due to the well established property of cascading failure propagation in the grid [11] and the small-world properties of power grids [13], a few initial points of failure due to a hurricane quickly cause network instability in a region potentially causing millions of people to suffer the effects of brownouts or blackouts. Due to built in failsafes and resilience mechanisms in the grid, the effect of failure of a component is often localized to a region depending on the severity of the fault or failure.

In the context of a hurricane, the grid component damage and failure is often progressive along the spatial trajectory of the hurricane. A hurricane exhibits multiple phases of varied intensity along its path causing failures with different levels of severity at different regions in its path. We model the progression of this grid failure process as a temporal segmentation problem.

Modeling this failure process over time, across different regions (e.g. counties) affected by a hurricane, is essential for improving the resilience of critical infrastructure during future disasters.

We characterize the severity of this grid failure process by measuring the number of people in a hurricane affected region (a county in our case) without power over the entire time period of the hurricane. Two critical questions need to be answered for characterization of this process:

- How can we characterize the different phases of a hurricane as a function of severity of the damage to critical infrastructure like the power grid using highly dynamic , sparse customer power loss data?
- Which locations (counties) are most important for characterizing each phase of failure?

Our main goal is to help domain experts answer the aforementioned questions.

Notation: We assume we are given a set of time series $X = \{x_1, x_2, ..., x_n\}$, where each time series $x_i = [x_i(t_1), x_i(t_2), \ldots, x_i(t_m)]$, and $x_i(t_j)$ represents the value at time stamp t_j for the i_{th} time series. We also assume there is an underlying graph structure G that captures the relation among these time series $\{x_i\}$, and we are given the Laplacian matrix L of G. For example, in critical infrastructure systems, the number of electric outages in all counties form a set of time series, and the relation among these counties can be based on the geographical proximity: two counties that are adjacent to each other are connected to each other. We use $x_i(t_a : t_b)$ to denote $[x_i(t_a), x_i(t_{a+1}), ..., x_i(t_b)]$, and $X(t_a : t_b) = [x_1(t_a : t_b), x_2(t_a : t_b), ..., x_n(t_a : t_b)]$.

Our algorithm CUT-n-REVEAL contains two parts: detecting a good segmentation of the outage data to capture the main changes; and finding the corresponding explanations for the segmentation. With this knowledge of the segmentation and the explanations for each segment, the expert has a holistic picture of the different phases of the failure process as well as the specific time series that contributed significantly to each phase change. We describe each of these tasks and our solutions in detail next.

3 FINDING SEGMENTATIONS

In this section we focus on the segmentation problem. We first describe the formal problem, and then give the overview and details of our method.

3.1 Our Problem

We first need to define a segmentation:

Definition 3.1 (Segmentation S). A segmentation of X contains a set of distinct time cut points $S = \{c_1, c_2, ..., c_k\}$, where $c_i \in \{t_1, t_2, \ldots, t_m\}$.

The cut points of S naturally divide the time period into a set of time segments. We denote such a time segment as $s_i = [c_{i-1}, c_i)$ with $c_0 = t_1$, and $c_{k+1} = t_m$. Our problem is:

Problem 1. ***Given*** *a set of time series X, the Laplacian matrix L of the underlying network, and a number k,* ***find*** *the k-segmentation of S that captures the main pattern changes in X.*

3.2 Overview of our approach

We develop a model to provide simple segmentations of the continuously changing grid failure process and interpretable explanations of the segmentations. We need to isolate temporal sequences into discrete segments such that the properties of the failure process in each segment differ from neighboring segments. The process of manually or algorithmically picking reasonable segments is non-trivial as segments that are too small fail to capture significant properties of the failure process while picking segments that are too large although capturing all failure process characteristics, do not highlight the differences between the various phases of the process. Since this process is highly dynamic and the failure dataset is highly sparse in nature, methods based on capturing long-term correlation [10] or invariant learning [20] from the data will be unable to perform adequately.

Owing to the sparse setting at hand and the lack of any long term correlations, we decided to adopt a latent factor modeling approach which has been shown to be effective in sparse settings like recommendation systems [15, 21]. The proposed segmentation procedure learns a latent representation for each time step using the data for dynamic power loss across affected counties. The results of the segmentation procedure are designed to be sparse in nature while managing to capture all the major trends in the process. We propose to use a data driven methodology (instead of setting a constant segment size) to automatically derive a set of appropriate time segments. Through the segmentation model, we wish to group time steps with similar patterns into contiguous time segments. To this end, we propose to represent each time step in latent space so as to capture its temporal relationship with the other time steps, and merge similar time steps into temporally contiguous segments. This methodology allows us to obtain segments where time steps in a segment are related to their past and future time steps in a similar way, allowing for simpler explanations later about pattern changes.

3.3 Details

In line with the overarching goal of discovering the different phases of the failure process in the power grid during natural disasters, we consider these phases or segments as a collection S of disjoint sets c_i. Each set c_i contains contiguous time steps that belong to segment i. We wish to discover a collection $S = \{c1, ..., c_k\}$ that minimizes similarity between any two neighboring sets c_i, c_j. Two sets c_i, c_j are said to be neighboring sets if $c_i = \{t_l, .., t_{l+\Delta l}\}$ and c_j contains t_{l-1} or $t_{l+\Delta l+1}$.

By doing so, each segment c_i would capture a different pattern from its neighboring segments (c_{i-1}, c_{i+1}), thus the segmentation S captures pattern changes in the time series. We employ the normalized cut framework which has been shown to work well in subspace clustering and segmentation tasks [29].

The normalized cut algorithm generates segments such that the similarity of time steps within each segment are maximized while the similarity between time steps in different segments is minimized.

The question of how to represent each time step, for effective similarity calculation between time steps still remains. One straightforward way is to use the data values at t_i and t_j across all time series and check how similar the values are. However, in this case, much of the information relating to the continuous evolution of the failure process will be lost.

In an effort to find a more principled approach to capture the similarity between different time steps in the failure process, we adopt the formulation provided by Tierney et al. [30] for video scene segmentation for our purposes of modeling the hurricane failure process. The model represents each time step in the data X, using a latent representation as a function of other important time steps. It is through this latent representation V that we attempt to capture the dynamics in the data X.

$$\begin{aligned} \min_{V} \ & \frac{1}{2}||X - XV||_F^2 + \lambda_1||V||_1 + \lambda_2||VR||_{1,2} \\ \textbf{s.t.} \ & \textbf{diag}(V) = 0 \end{aligned} \tag{1}$$

where V is an m by m matrix, and the i^{th} column represents the latent representation of time step i in terms of all the other time steps.

The first term in Eq. 1 calculates the reconstruction error between X and XV while, the second term introduces sparsity into the latent representation, enforcing that, each time step be explained as a function of a small subset of other important time steps. Finally in the third term, R is an m by $m-1$ lower triangular matrix with -1 on the diagonal and 1 on the second diagonal.

$$R = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{bmatrix} \tag{2}$$

The term VR calculates the difference of each time step with its previous time step in the latent V space. This term essentially serves as a smoothness constraint penalizing the dissimilarity of neighboring time steps. The $l_{1,2}$ norm term forces whole column similarity between two columns of V, i.e. between neighboring time steps in V as opposed to just element-wise similarity in the case of a simpler l_1 norm on VR.

The solution to equation 1 can be obtained by applying the alternating direction method of multipliers (ADMM) [4]. Using this approach does not guarantee convergence but in our experiments, the algorithm always converged. To separate each term in equation 1, we assign $K = V$ and $U = KR$.

The Lagrangian formulation is given by equation 3

$$\begin{aligned} \mathcal{L}(V, K, U) = & \frac{1}{2}||X - XK||_F^2 + \lambda_1||V||_1 + \lambda_2||U||_{1,2} \\ & + \langle G, V - K\rangle + \frac{\beta_1}{2}||V - K||_F^2 + \langle F, U - KR\rangle \\ & + \frac{\beta_2}{2}||U - KR||_F^2 \end{aligned} \tag{3}$$

Solving equation 3 yields a temporal weight matrix $V \in R^{m \times m}$ from which we derive an affinity W. The affinity matrix is then segmented using the normalized cuts procedure as mentioned previously to obtain the set of segments S.

4 FINDING EXPLANATIONS

Despite the sparsity of the segmentation procedure in the previous section, it is often not possible to identify the cause for each segment due to many simultaneously changing time series through out the failure process. In this scenario, it is beneficial for a domain expert to know the subset of "culprit" time series that were influential in the model selecting a particular segment.

4.1 Our Problem

Assuming we are given a segmentation S of X, containing a set of time segments and the corresponding cut points $\{c_i\}$, which counties/time series are the most important for characterizing the different patterns in the segments? How to design explanations of the segmentations in an intuitive, easy to understand way? The cut points in the segmentation and our latent V matrix (which instead only captures the temporal relation among *time stamps*) do not provide answers to this question. A desired explanation of the segmentation should be simple yet effective enough that they give direct guidance to prevent or curtail the effect of the failure of critical infrastructure in future disasters.

We provide answers to these questions by introducing an explanation vector e_i for each cut point c_i in the segmentation. Each e_i is an n by 1 vector, where the j_{th} value represents the importance of the j_{th} time series/counties in explaining the cut point. Intuitively, if time series x_j shows very different patterns before and after the cut point c_i, we consider it important in explaining why c_i is a good cut point. On the other hand, if time series remains constant/unchanged across c_i, it does not provide useful information in terms of the cut point c_i and should have low values in e_i. In the hurricane outage data where there are hundreds of time series/counties, such explanation vectors are able to highlight the culprit time series/counties where major changes happen at the cut point. This explanation also serves as a guide for future resource allocation policies of maintenance and emergency personnel.

Definition 4.1 (Cut point explanations E). $E = \{e_1, e_2, ..., e_k\}$, where e_i is an n by 1 non-negative explanation vector. $||e_i||_1 = 1$ and e_{ij} represents the importance of time series j for explaining the cut point c_i.

Now we give the problem we solve.

PROBLEM 2. ***Given*** *a set of time series X, the Laplacian matrix L of the underlying network, a number k, and the k-segmentation of S,* ***find*** *the associated explanations E, that capture the main pattern changes in X.*

4.2 Overview of our approach

Existing time series segmentation algorithms do not provide any explanation of the result in an automatic principled way. Only recently there has been a push toward making complex machine learning model outputs quantifiable, explainable and simple [25]. To design good explanations specifically for hurricane outage data, we consider the characteristics of the data, as well as the requirements from the domain experts. We formulate an optimization problem that automatically learns explanations considering the underlying geographical relation between counties, revealing a small number of truly important counties as the culprits for the domain experts.

4.3 Details

We want to design an optimization problem that automatically finds good $\{e_i\}$. Assume that we have a function $d(S, i)$, which takes a segmentation S and a cut point index i as inputs, and returns an n by 1 vector which captures the difference of each time series before and after the i_{th} cut point c_i in S. We want e_i to give higher weights on time series with higher $d(S, i)_j$ values (therefore higher difference across cut point c_i). A straight forward way is to maximize the weighted sum in the following way.

$$\begin{aligned} \arg\max_E \; & \sum_{i=1}^{k} e_i^T d(S,i) \\ \text{subject to} \quad & 0 \le e_{ij} \le 1, \\ & ||e_i||_1 = 1 \end{aligned} \tag{4}$$

The above formulation guarantees that we give higher importance to time series/counties with larger difference across the corresponding cut points. However, it treats each county as independent and ignores their geographical relation (some counties are close in distance, and some are far away). Such a geographical relation is important to capture because a hurricane trajectory is continuous and it usually hits counties that are close to each other at the same time. Hence, the importance of counties should be geographically smooth in the sense that adjacent counties should have similar importance. Another drawback of the above formulation is that it does not correctly reflect our requirement of a 'simple' explanation: we only need a few culprit counties and want to avoid having high importance for too many counties. Due to these considerations, we improve the above formulation by adding a geographical and a sparsity constraints on e_i. The final optimization problem we solve is shown below.

Given: A set of time series X,L, a segmentation S, α, λ.
Find: $E = \{e_i\}$ such that

$$\begin{aligned} \arg\max_E \; & \sum_{i=1}^{k} [e_i^T d(S,i) - \alpha e_i^T L e_i] - \lambda \sum_{i=1}^{k} ||e_i||_1 \\ \text{subject to} \quad & 0 \le e_{ij} \le 1, \\ & ||e_i||_1 = 1 \end{aligned} \tag{5}$$

The geographical smoothness is introduced in the second term using the Laplacian matrix L (obtained from the underlying network behind the counties). This term basically minimizes the difference of e_i for adjacent counties. The third term is an $L1$ norm regularization on e_i, which introduces sparsity in e_i. This leads to the simplicity in our explanations: only a few important counties will have non-zero values in e_i to explain the c_i), making the results much simpler to interpret.

One question remains unsolved is how should we design the distance function $d(S, i)$ to capture the difference of time series across a cut point? Our idea is to look at a time window before the cut point c_i and a time window after c_i, and calculate the difference of these two time windows as the difference of the time series across c_i. Assume that w_{ij}^- represents the sub-sequence of x_j in the time window before c_i, and w_{ij}^+ represents the sub-sequence in the time window after c_i. We then calculate the difference of w_{ij}^- and w_{ij}^+ using simple, standard time series features: the mean value (f_1), the standard deviation (f_2), the maximum value (f_3) and the minimum value (f_4).

$$D(S,i)_j = \frac{1}{4} \sum_{z=1}^{4} |f_z(w_{ij}^-) - f_z(w_{ij}^+)| \tag{6}$$

Note that $|f_z(w_{ij}^-) - f_z(w_{ij}^+)|$ are usually in different magnitudes for different features, hence as a preprocessing step which we do not elaborate in the equation, we perform a min-max normalization of $|f_z(w_{ij}^-) - f_z(w_{ij}^+)|$ across all time series to make them in the same scale. As both w_{ij}^- and w_{ij}^+ are of a short length (a deliberate setting since the pattern changes that justify the choice of a particular cut points usually lie in the local area), these simple features are enough to capture the main pattern difference.

Finally, to solve Eq. 6, we optimize each e_i separately. For each e_i, the optimization can be re-written as a Quadratic Programming problem in the following way.

$$\begin{aligned} \arg\min_{e_i} \quad & \alpha e_i^T L e_i - [d(S,i)^T - \lambda i^T] e_i \\ \text{subject to} \quad & 0 \le e_{ij} \le 1, \\ & ||e_i||_1 = 1 \end{aligned} \tag{7}$$

The QP problem is well studied in the literature, and it is NP-hard in its general form. In our case, where the QP is convex to e_i, it can be solved in polynomial time using an Interior Point method [33], and we use the existing Matlab function (quadprog) to solve the problem.

5 EMPIRICAL STUDY

We implement CUT-n-REVEAL in Python and Matlab (we will release the code for research purposes). Our experiments were conducted on a 4 Xeon E7-4850 CPU with 512 GB of 1066Mhz main memory.

5.1 Set-up

Dataset. We collect datasets from different domains with the ground truth segmentations to quantitatively evaluate our performance, and we run CUT-n-REVEAL on three hurricane outage datasets to show our case study results. For efficiency purposes, we perform a standard rolling average as a preprocessing step to all the data. The final statistics of the datasets are shown in Tab. 1.

ChickenDance: A "chicken" dance motion is recorded as a sequence of 4-dimensional data points. This was extracted by Matsubara et.al [18] and is originally from CMU motion capture database[2]. We have the ground-truth segmentation here based on motions in the dances.

WalkJog: We use a dataset adapted from the REALDISP Activity Recognition Dataset [1, 9], where motions of walking, jogging and running are recorded by sensors. We use a sub-dataset with the walking motion and the jogging motion (as these motions show obvious different patterns), and we aim to detect when a different motion happens.

NILM: Non Intrusive Load Monitoring dataset. This dataset consists of real power measurements for various household appliances like lamps, laptops, and refrigerators, recorded through the use of MAU (Measurement and Actuation Units) connected between the device and the wall-socket. A detailed account of this dataset and the data acquisition methodology has been presented in [24]. We use a twenty-four hour snapshot of the NILM data to evaluate our performance, and use the time when a device switch states as the ground truth segmentation.

Hurricane Outage data: ORNL has developed several grid situational awareness products over the last decade such as VERDE, EARSS and EAGLE-I[3] for different stakeholders like DOE and FEMA, primarily for emergency management. For example, the National Outage Map within EAGLE-I collects distribution outage data of all the customers from utility websites every 15 minutes. Due to the recent coverage expansion (with more utilities exposing data from their Outage Management Systems), in this paper, we consider the more recent hurricane outage data namely for Matthew, Harvey and Irma since it covers nearly 90% of the population in the hurricane affected areas.

Baselines To the best of our knowledge, there is no algorithm that finds explanations for segmentations in the way we do. In the experiments, we mainly select state-of-the-art multi-variate time series segmentation algorithms to compare against.

Autoplait [18] is a Hidden Markov Model based algorithm that discovers the different regimes in co-evolving time series. Each regime can be thought of as the segments for our problem.

TICC [10] is a subsequence clustering algorithm for multivariate time series to discover repeated patterns. It clusters time stamps into segments that can be well interpreted by the same model. Each cluster of the time stamps becomes the segments for our problem.

Dynammo [17] reconstructs data values in time series by discovering latent variables and their dynamics, and then uses the spikes of the reconstruction errors to find cut points for the data.

Table 1: Datasets used.

Dataset	#Timestamps	#Time series
NILM	721	5
ChickenDance	322	4
WalkJog	500	2
Harvey	264	250
Irma	169	271
Matthew	252	369

5.2 Quantitative Evaluation

We compare CUT-n-REVEAL performance with the baselines on the datasets with ground truth segmentations: *NILM, ChickenDance* and *WalkJog*. We evaluate the detected cut points by calculating the F1 score based on the ground truth cut points (similarly as in [18]). Higher F1 scores show better segmentation precisions. We show the results in Tab. 2.

Table 2: Evaluation on Ground Truth Datasets

Method / Dataset	CUT-n-REVEAL	AutoPlait	TICC	Dynammo
NILM	0.8	0.4	**0.81**	0.75
ChickenDance	**0.8**	0.73	0.5	0.57
WalkJog	**1.0**	0.0	0.25	0.0

As shown in the table, except for the *NILM* data, CUT-n-REVEAL significantly outperforms the baselines (achieving the best F1 scores) in most of the cases. Even in *NILM*, our performance is comparable to TICC's (a difference of only 0.01). This showcases the effectiveness of CUT-n-REVEAL at identifying and extracting different patterns in time series. In the following, we visualize our segmentations and show that they capture the main pattern changes in the data.

We show our segmentation result for *ChickenDance* in Fig. 1. CUT-n-REVEAL is able to isolate most of the different data trends successfully. It detects precisely 6 cut points from the 7 ground truth cuts, with two false positive cuts at time steps 72 and 234, which are very close to ground truth cuts. In Fig. 2, for *WalkJog*, we see that CUT-n-REVEAL correctly separates the sequences of data generated due to walking from those due to jogging. The cut point discovered by our method lies in a 1% cut point location tolerance window with respect to the ground truth cut point. Finally for the *NILM*

[2] http://mocap.cs.cmu.edu
[3] Outage Data source: EAGLE-I https://eagle-i.doe.gov/

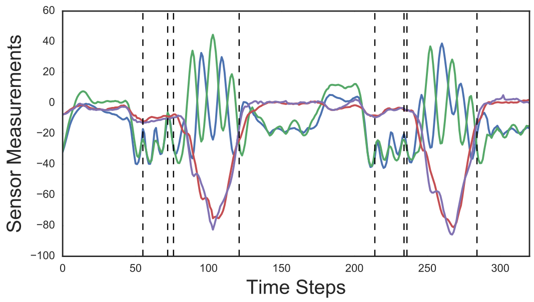


Figure 1: The discovered segmentation results for *ChickenDance*.

dataset, the results (Figure 3) are able to identify accurately (F1 score=0.8) the different residential daily usage patterns (segments 1 - 9 enumerated from left to right of the figure).

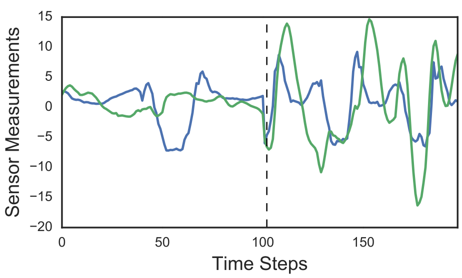


Figure 2: The discovered segmentation (vertical black dashed line) results for the *WalkJog* dataset.

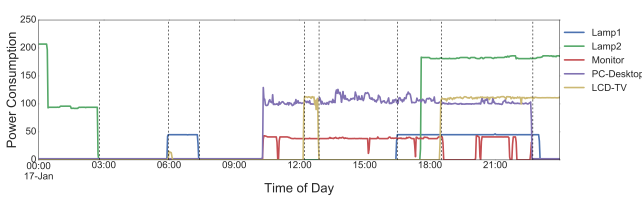


Figure 3: The discovered segmentation (vertical black dashed line) results for the *NILM* dataset.

5.3 Case Studies: NILM

Monitoring the energy consumption pattern of a home is essential for undersatnding patterns of power wastage and overuse. Hence it is interesting to study in-depth the patterns as shown by our segmentation for the *NILM* dataset. In Figure 3, each segment captures a state change of one or multiple devices. Segments 1 and 2 represent a possible nightly routine with all electronic devices except possibly a night lamp (Lamp2) turned off. Segment 4 represents the power consumption during the first half of the work day and segment 5 captures the event of the television being turned on, possibly during a lunch break. Segment 6 could be equated with segment 4 representing the second half of the work day. Segment 7 represents the beginning of the evening routine. This segment is also the period of peak demand for the day. If a user wished to lower their energy bill, it is apparent that, the energy consumption in this segment would have to be reduced or distributed to neighboring segments. Segment 8 and 9 can be considered evening recreational activity. The segmentation algorithm thus allows a resident to gain an abstracted view of their energy usage patterns as a function of their daily activities with the devices in each segment highlighted. In order to further our performance evaluation, To understand if our results are robust to noise (common in such datasets), we also included another device (a refrigerator) that consumes significant energy relative to the other appliances, and has a consistent cyclic daily usage pattern. The usage pattern with segmentation on the new dataset is shown in Figure 4. Interestingly CUT-n-REVEAL is quite robust in the presence of such devices. It is still able to effectively capture the device state changes of all appliances in the household like before; the difference being the additional segments during periods of isolated operation of the refrigerator.

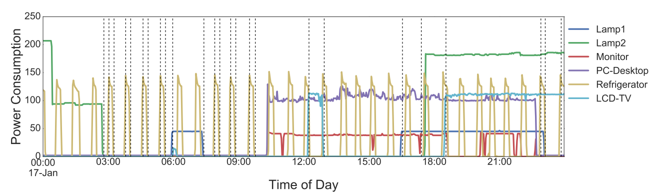


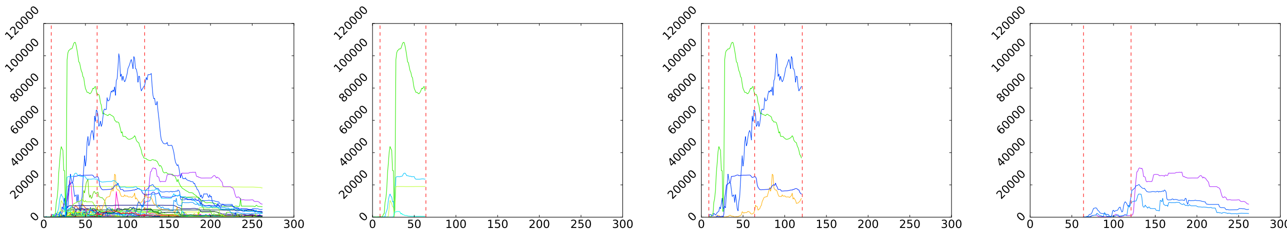
Figure 4: *NILM* with noisy cyclic patterns

5.4 Case Studies: Hurricanes

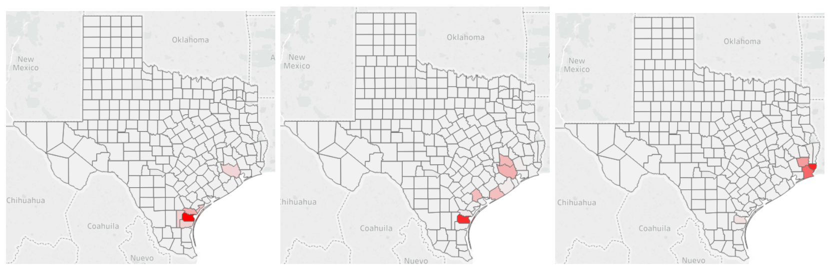
We run CUT-n-REVEAL on outage data from three recent hurricanes. We show that, thanks to CUT-n-REVEAL, we can find the segmentations which capture important outage pattern changes, and our explanations correctly identify the culprit counties where major events happen.

Hurricane Harvey. We show our segmentation and explanation results in Fig 5. The segmentation and all the time series are shown in Fig. 5(a). In Fig. 5(b)(c)(d), we visualize for each cut point the most important time series whose e_i values sum over 0.8 (their importance in explaining the corresponding cut point is over 80%). In Fig. 5(e)(f)(g), we visualize the entire e_i vector on a map, where the color of a county shows the importance of that county in explaining the cut point.

We detect three cut points for Harvey (see Fig. 5(a)). The first cut point captures the date (Aug 25) when the hurricane strengthened before its landfall. Soon after this cut point, the number of outages starts to rise in several counties. In Fig. 5(b)(e), we observe that CUT-n-REVEAL correctly highlights the counties with the steepest rise for the first cut point (Nueces, San Patricio and Aransas), and all of them are near the Harvey's landfall area. For the second cut point (around Aug 27), our explanations capture two different patterns: the number of outages in the Nueces county (green line in Fig. 5(c) and the red county in Fig. 5) experience major decrease, while the number of outages in the Victoria, Matagorda and Montegomery counties start to rise. This cut point correctly shows that the impact of the hurricane is moving, and our explanations identifies the



(a) Harvey segmentation (b) Important time series for c_1 (c) Important time series for c_2 (d) Important time series for c_3



(e) Explanation e_1 (f) Explanation e_2 (g) Explanation e_3

Figure 5: Segmentation and the corresponding explanations for Harvey. (a) The segmentation and all the time series. (b)(c)(d) The most important time series (that contribute over 80% importance in e_i) for each of the cut point. (e)(f)(g) e_i visualizations for each cut point. Counties with higher e_i values are more important for the cut point, and are mark with a color closer to red.

direction of the move (which is consistent with the trajectory of Harvey). Finally for the last cut point (around Aug. 30), while the outages of many counties are decreasing, our algorithm correctly detects a small set of counties (Orange, Jefferson and Hardin) that experience a sudden outage increase (Fig. 5(d)(g)). The main reason for this increase is due to the rising water of the Neches River, which causes the city to lose service from its major pump stations. Note that this information is deeply buried in the entire time series, it is hard to directly observe in Fig. 5(a) with hundreds of time series why such a cut point make sense. Through our explanations, we perfectly isolate these culprit counties in Fig. 5(d)(g).

Hurricane Irma. We make similar plots for Irma in Fig. 6. The first cut point (around Sep 10) again captures the date of the hurricane's landfall. Interestingly, the second and third cut point of the segmentation are very close to each other, and by directly reading the raw data with hundreds of time series in Fig. 6(a), it's impossible to make sense of these two cut points. On the other hand, the explanations learned by CUT-n-REVEAL precisely show why these two cut points are detected: they are correspondent to outage increases in different locations that are far away from each other. The second cut point shows the outage increase in the Pinellas, Duval and St. Johns counties, while the third cut point captures the outage outbreak in the DeKalb, Fulton and Gwinnett counties (the highlighted area in the top of Fig. 6(h)). Finally, the last cut point (around Sep 12) correctly captures the date when hurricane Irma weakened into a Category 2 storm, and the number of outages in many counties decrease.

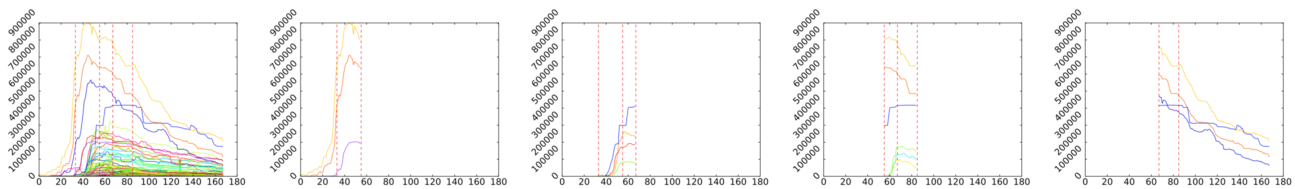
Hurricane Matthew. In Fig. 7, we show our segmentation and explanations in similar ways for the hurricane Matthew. The first cut point (around Oct 2) corresponds to the landfall of Matthew and a set of counties in the east coast (Brevard, Palm Beach, etc.) start to have increasing power outages. The second cut point (around Oct 4) and the corresponding explanations show the moving direction of the hurricane. Counties like Duval, Chatham, Horry and Charleston, which are to the North of the landfall location start to experience power outages (the Chatham county experience a peak of outage at the cut point). Soon after Oct 4, in Fig. 7(h), we observe that the Horry county (highlighted with the red color), which is influenced in the previous cut point, has now become severely affected and the influence has spread to neary counties as well. This can be also observed in Fig. 7(d) where a set of time series suddenly increase. Finally, the last cut point (around Oct 9) shows that the power outage impact of Matthew on these previously affected counties has mostly abated.

6 RELATED WORK

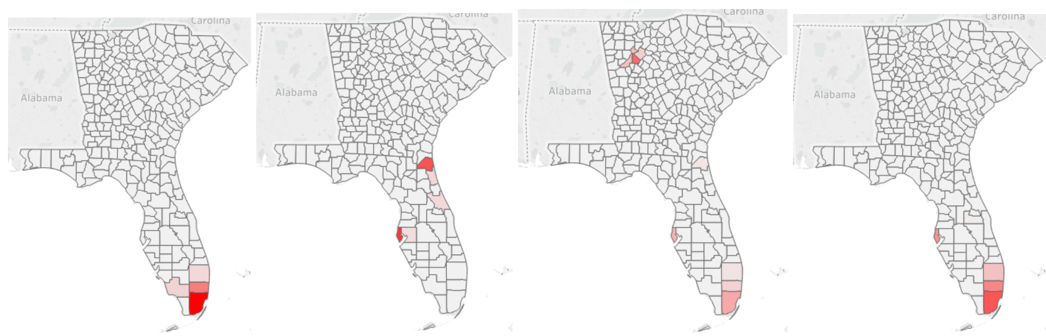
We review the most related work in segmentation below.

Time series segmentation. There has been a lot of work on time series, such as modeling co-evolving time series using multi-level HMMs [18], discovering patterns in data streams [26], developing online algorithms for frequent sequence mining [19], time series segmentation [17, 28]. Change point detection has also been a popular topic in the climate sciences as well [23].

Characterizing the dynamics of natural disasters like hurricanes lends itself naturally to a change point detection approach but there has been little work conducted in this regard. [34] proposes a hierarchical Bayesian framework for detecting shifts in annual hurricane counts while [27] introduces a Bayesian change point

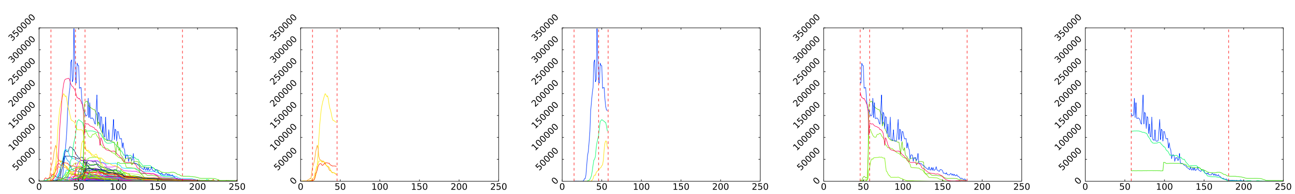


(a) Irma segmentation (b) Important time series for c_1 (c) Important time series for c_2 (d) Important time series for c_3 (e) Important time series for c_4

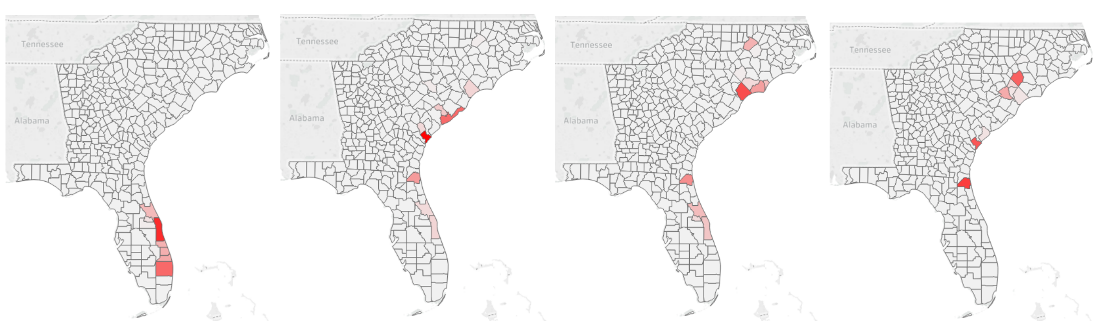


(f) Explanation e_1 (g) Explanation e_2 (h) Explanation e_3 (i) Explanation e_4

Figure 6: Segmentation and the corresponding explanations for Irma. See detail discussions in Sec. 5.4.



(a) Matthew segmentation (b) Important time series for c_1 (c) Important time series for c_2 (d) Important time series for c_3 (e) Important time series for c_4



(f) Explanation e_1 (g) Explanation e_2 (h) Explanation e_3 (i) Explanation e_4

Figure 7: Segmentation and the corresponding explanations for Matthew. See detail discussions in Sec. 5.4.

detection algorithm to detect changes in temperature using climate data records. The other line of work on modeling failure cascading on CIS [5] do not explicitly segment the time series. Despite the extensive research conducted in time series segmentation and change point detection, we have found that there exists little prior work in leveraging them to characterize the dynamic effect natural disasters have on critical infrastructure systems.

Another line of related work in time series are the subspace clustering based techniques. Many applications in multivariate time series analysis exist wherein the temporal data is drawn from multiple spaces and hence exhibits multi-segment behavior. It is often useful to develop techniques to represent the data in a subspace to capture richer temporal relationships and apply clustering to explicitly demarcate these multiple segments. This approach called subspace clustering has been applied to video and image segmentation [30, 32], image compression [14] and spatio-temporal action segmentation [6, 16]. A comprehensive review about the different types of subspace clustering methods is provided by [31]. There has been extensive work in subspace clustering in data mining [22] but to the best of our knowledge, it has not been applied on the hurricane outages data for finding the temporal relation among time steps. Further, these subspace clustering techniques also do not provide explanations of the results.

Simple Interpretable Models. There has recently been a push towards quantifying model uncertainty [8] and making machine learning model outputs quantifiable, explainable and simple [25]. These models and their explanations are specific to the underlying machine learning models and cannot be applied for our segmentation problem. We find that the temporal segmentations are inherently unsupervised and the intuition behind the segments might not be readily apparent or explainable in certain applications. To the best of our knowledge, our explanation optimization problem is the first attempt toward designing simple explanations for time series segmentations.

7 CONCLUSIONS

In this paper, we have developed a combined framework for providing simple interpretable explanations for failure processes like critical infrastructure outages. We evaluated the performance of our methodology against state of the art segmentation and time series clustering procedures on open ground truth datasets. We have also conducted extensive analysis on the failure of the power grid during three hurricane events along with conducting a case study on the applicability of temporal segmentation to understanding residential energy usage patterns. There are many avenues for future work. Methodologically, we can study performing a joint learning of segmentations and explanations to leverage both spatial and temporal information simultaneously. We are also exploring integrating CUT-n-REVEAL with existing analysis tools, such as the URBANNET toolkit [5] in use at national labs and power utilities.

REFERENCES

[1] Oresti Banos, Mate Attila Toth, Miguel Damas, Hector Pomares, and Ignacio Rojas. 2014. Dealing with the effects of sensor displacement in wearable activity recognition. *Sensors* 14, 6 (2014), 9995–10023.

[2] Alan M Barker, Eva B Freer, Olufemi A Omitaomu, Steven J Fernandez, Supriya Chinthavali, and Jeffrey B Kodysh. 2013. Automating natural disaster impact analysis: An open resource to visually estimate a hurricane's impact on the electric grid. In *Southeastcon, 2013 Proceedings of IEEE*. IEEE, 1–3.

[3] Eric S Blake, Todd B Kimberlain, Robert J Berg, John P Cangialosi, and John L Beven Ii. 2013. Tropical cyclone report: Hurricane sandy. *National Hurricane Center* 12 (2013), 1–10.

[4] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning* 3, 1 (2011), 1–122.

[5] Liangzhe Chen, Xinfeng Xu, Sangkeun Lee, Sisi Duan, Alfonso G Tarditi, Supriya Chinthavali, and B Aditya Prakash. 2017. HotSpots: Failure Cascades on Heterogeneous Critical Infrastructure Networks. In *CIKM*. ACM, 1599–1607.

[6] Ehsan Elhamifar and René Vidal. 2009. Sparse subspace clustering. In *CVPR*. IEEE, 2790–2797.

[7] Rozhin Eskandarpour and Amin Khodaei. 2017. Machine learning based power grid outage prediction in response to extreme events. *IEEE Transactions on Power Systems* 32, 4 (2017), 3315–3316.

[8] Zoubin Ghahramani. 2015. Probabilistic machine learning and artificial intelligence. *Nature* 521, 7553 (2015), 452.

[9] Shaghayegh Gharghabi, Yifei Ding, Chin-Chia Michael Yeh, Kaveh Kamgar, Liudmila Ulanova, and Eamonn Keogh. 2017. Matrix Profile VIII: Domain Agnostic Online Semantic Segmentation at Superhuman Performance Levels. In *ICDM*. IEEE, 117–126.

[10] David Hallac, Sagar Vare, Stephen Boyd, and Jure Leskovec. 2017. Toeplitz inverse covariance-based clustering of multivariate time series data. In *KDD*. ACM, 215–223.

[11] Paul Hines, Karthikeyan Balasubramaniam, and Eduardo Cotilla Sanchez. 2009. Cascading failures in power grids. *IEEE Potentials* 28, 5 (2009).

[12] Paul DH Hines, Ian Dobson, and Pooya Rezaei. 2017. Cascading power outages propagate locally in an influence graph that is not the actual grid topology. *IEEE Transactions on Power Systems* 32, 2 (2017), 958–967.

[13] Åke J Holmgren. 2006. Using graph models to analyze the vulnerability of electric power networks. *Risk analysis* 26, 4 (2006), 955–969.

[14] Wei Hong, John Wright, Kun Huang, and Yi Ma. 2006. Multiscale hybrid linear models for lossy image representation. *IEEE Transactions on Image Processing* 15, 12 (2006), 3655–3671.

[15] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009).

[16] Colin Lea, Austin Reiter, René Vidal, and Gregory D Hager. 2016. Segmental spatiotemporal cnns for fine-grained action segmentation. In *ECCV*. Springer, 36–52.

[17] Lei Li, James McCann, Nancy S Pollard, and Christos Faloutsos. 2009. Dynammo: Mining and summarization of coevolving sequences with missing values. In *KDD*. ACM, 507–516.

[18] Yasuko Matsubara, Yasushi Sakurai, and Christos Faloutsos. 2014. Autoplait: Automatic mining of co-evolving time sequences. In *SIGMOD*. ACM, 193–204.

[19] Abdullah Mueen and Eamonn Keogh. 2010. Online Discovery and Maintenance of Time Series Motifs *(KDD '10)*. 1089–1098.

[20] Nikhil Muralidhar, Chen Wang, Nathan Self, Marjan Momtazpour, Kiyoshi Nakayama, Ratnesh Sharma, and Naren Ramakrishnan. 2018. illiad: InteLLigent Invariant and Anomaly Detection in Cyber-Physical Systems. *ACM TIST* 9, 3 (2018), 35.

[21] Xia Ning and George Karypis. 2011. Slim: Sparse linear methods for top-n recommender systems. In *ICDM*. IEEE, 497–506.

[22] Lance Parsons, Ehtesham Haque, and Huan Liu. 2004. Subspace clustering for high dimensional data: a review. *Acm Sigkdd Explorations Newsletter* 6, 1 (2004), 90–105.

[23] Jaxk Reeves, Jien Chen, Xiaolan L Wang, Robert Lund, and Qi Qi Lu. 2007. A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology* 46, 6 (2007), 900–915.

[24] Andreas Reinhardt, Paul Baumann, Daniel Burgstahler, Matthias Hollick, Hristo Chonov, Marc Werner, and Ralf Steinmetz. 2012. On the accuracy of appliance identification based on distributed load metering data. In *SustainIT (2012)*. IEEE, 1–9.

[25] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. 2016. Why should i trust you?: Explaining the predictions of any classifier. In *KDD*. ACM, 1135–1144.

[26] Guy Rosman, Mikhail Volkov, Dan Feldman, John W Fisher III, and Daniela Rus. 2014. Coresets for k-segmentation of streaming data. In *NIPS*. 559–567.

[27] Eric Ruggieri. 2013. A Bayesian approach to detecting change points in climatic records. *International Journal of Climatology* 33, 2 (2013), 520–528.

[28] Allou Samé and Gérard Govaert. 2012. Online Time Series Segmentation Using Temporal Mixture Models and Bayesian Model Selection. *ICMLA* 1 (2012), 602–605.

[29] Jianbo Shi and Jitendra Malik. 2000. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence* 22, 8 (2000), 888–905.

[30] Stephen Tierney, Junbin Gao, and Yi Guo. 2014. Subspace clustering for sequential data. In *Proceedings of IEEE CVPR*. 1019–1026.

[31] René Vidal. 2011. Subspace clustering. *IEEE Signal Processing Magazine* 28, 2 (2011), 52–68.

[32] Allen Y Yang, John Wright, Yi Ma, and S Shankar Sastry. 2008. Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding* 110, 2 (2008), 212–225.

[33] Yinyu Ye and Edison Tse. 1989. An extension of Karmarkar's projective algorithm for convex quadratic programming. *Mathematical programming* 44, 1-3 (1989), 157–179.

[34] Xin Zhao and Pao-Shin Chu. 2006. Bayesian multiple changepoint analysis of hurricane activity in the eastern North Pacific: A Markov chain Monte Carlo approach. *Journal of climate* 19, 4 (2006), 564–578.