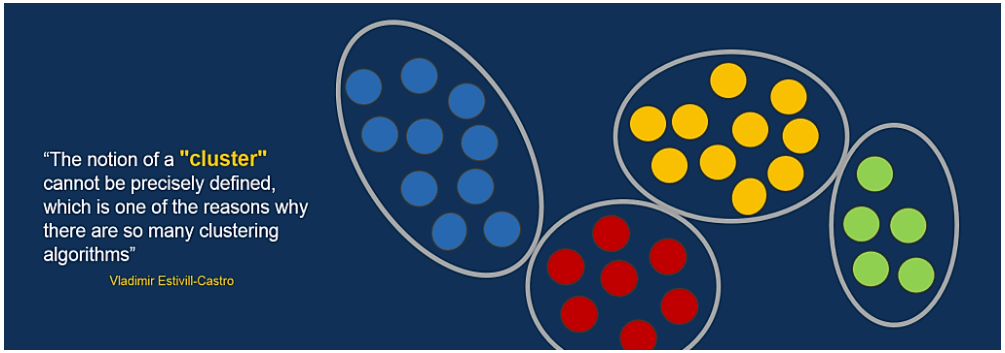
SAND2016-7809PE

Exceptional service in the national interest

Sandia National Laboratories

Pruning Dynamic Event Trees Using Density Peaks with Dynamic Time Warping



Nevin Martin
University of New Mexico

Supervised by
Matthew Denman
Org. 6231

6220/6230 summer student mini-symposium
August, 16th 2016, Sandia Nat'l Labs, Albuquerque NM

U.S. DEPARTMENT OF ENERGY | National Nuclear Security Administration | LDRD Laboratory Directed Research & Development

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Outline

Sandia National Laboratories

1. Pruning Dynamic Event Trees
2. Density Peaks Clustering Algorithm
3. Preliminary Results and Clustering for Regression
4. Discussion
5. Future Work

Static event trees can oversimplify severe accidents

Sandia National Laboratories

Initiating Event: Fire

Static

Assumes the firetruck will arrive in 30 minutes.

Dynamic

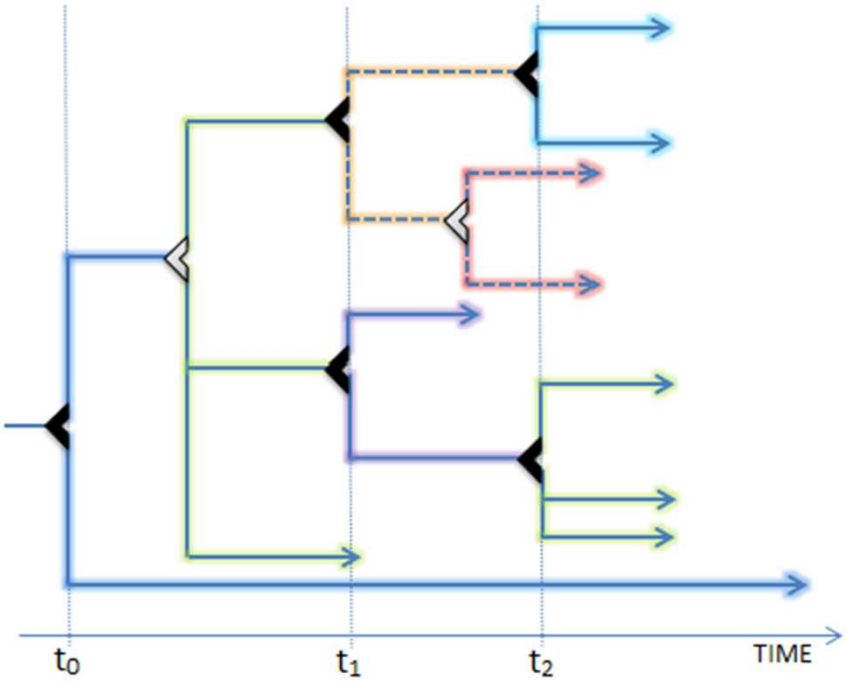
Models the firetruck driving to the house.

Accounts for possible traffic jams, car accidents, etc. that may effect arrival time.

When and Where to Prune?

The key choice that must be made is **when to stop the growth of a branch**.

Pruning can be performed at any branch point in a dynamic event tree. We propose a pruning algorithm that uses density peaks clustering with DTW.

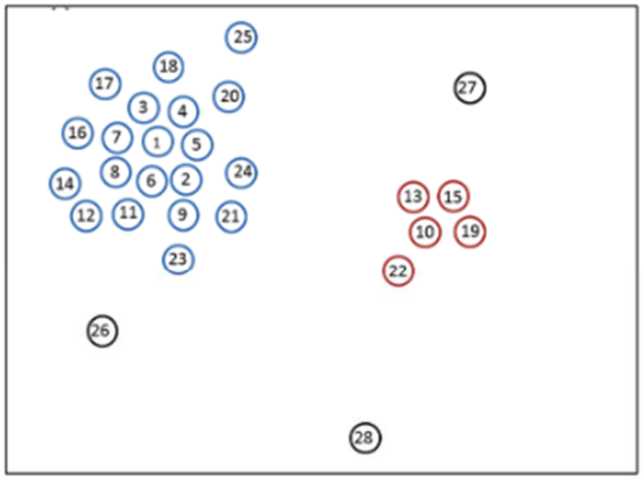


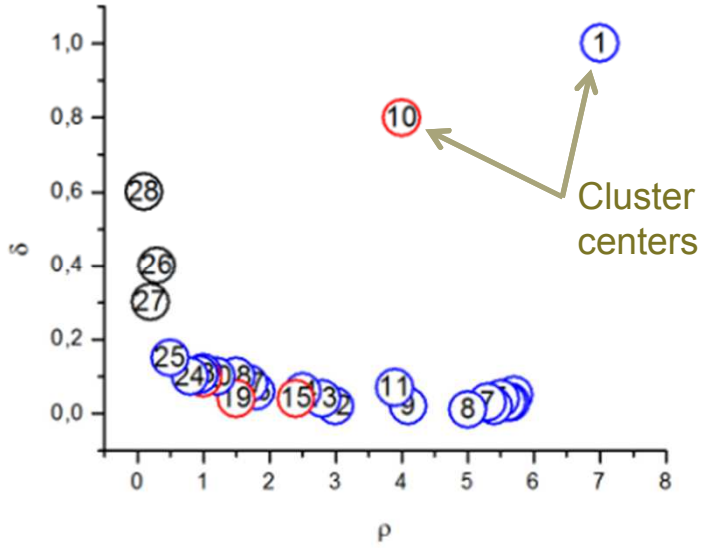
- Cluster centers will be determined at each branch point.
- Branches from clusters will be pruned.
- The number of branches pruned will be related to user-defined acceptable amount of fidelity reduction in the final tree.

Density Peaks Clustering Algorithm

Steps:

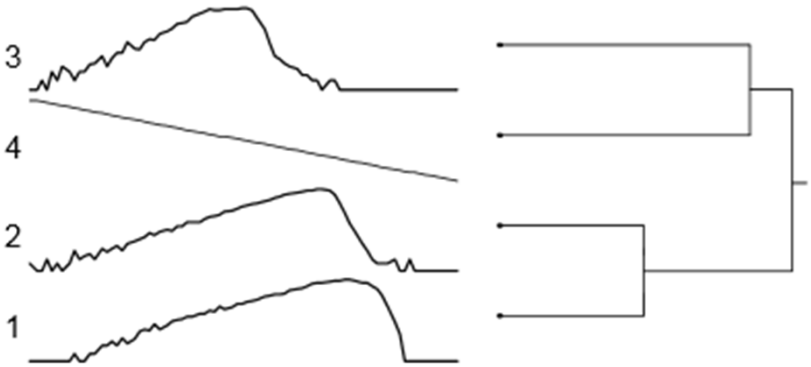
1) Create a pairwise distance matrix for two time series.
2) For each point in the matrix, calculate **the number of data points within a user-determined distance**, ρ_i.
3) Then, for each point, calculate **the distance of the closest data point of higher density**, δ_i .
4) Cluster centers are defined as those points that have the highest values of ρ_i x δ_i .



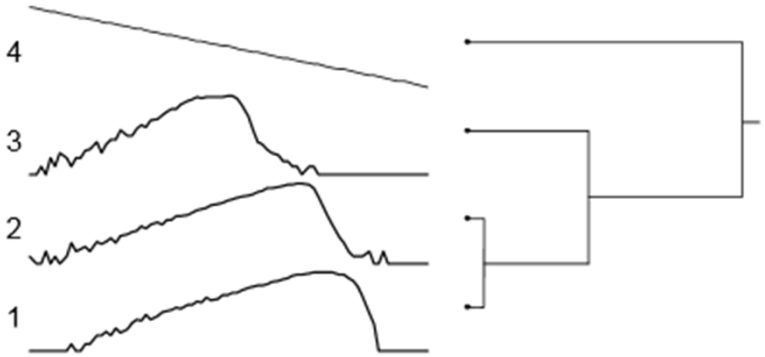


How to Create a Distance Matrix?

- Dynamic Time Warping (DTW) calculates the pairwise distance between all points in two time series.
- DTW is superior to the Euclidean distance as it can identify similarities in two time series even if there is a time lag.

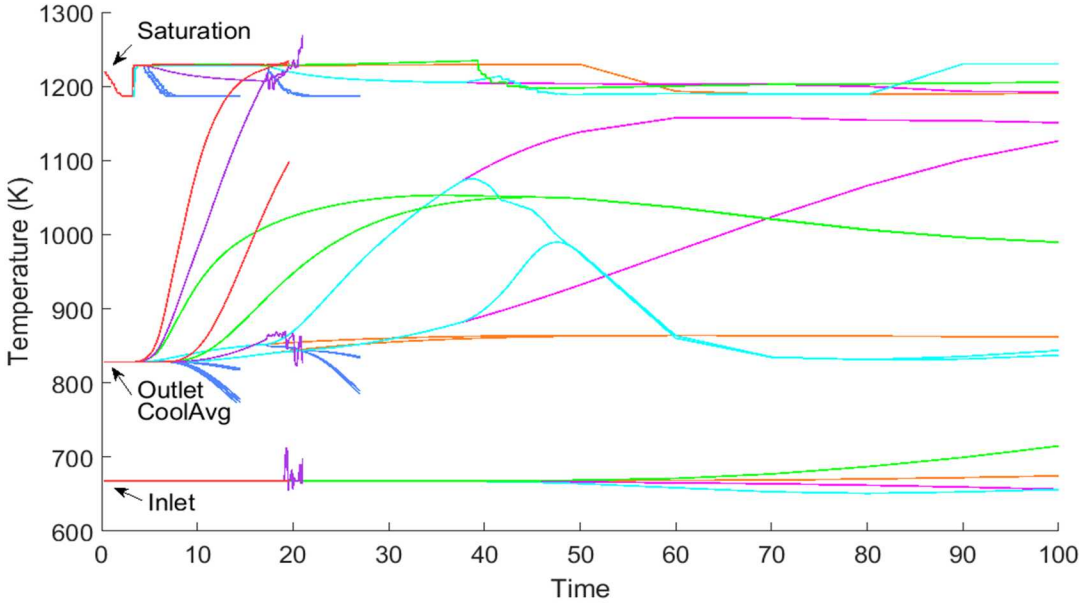


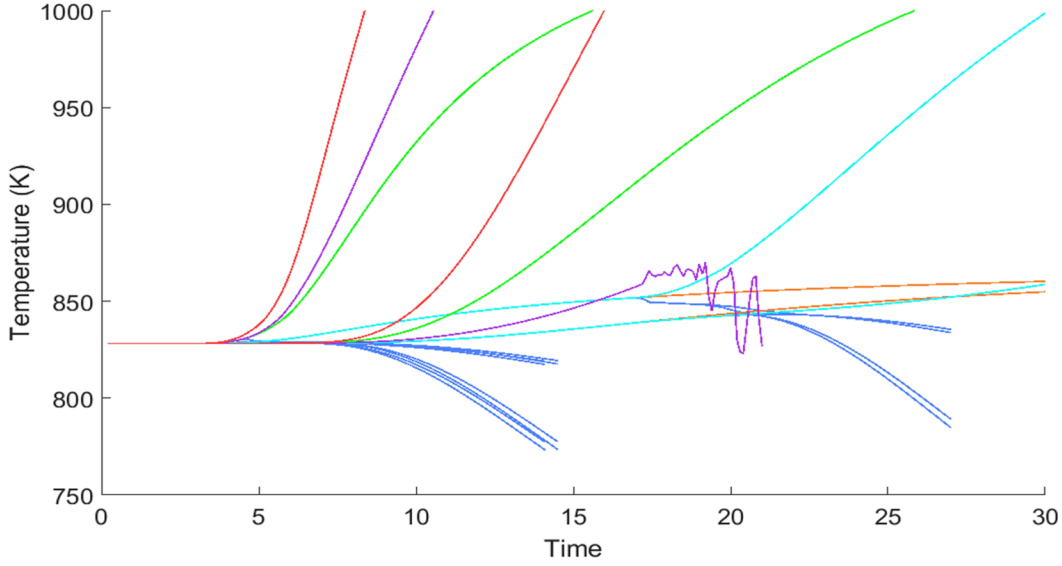
Clustering using Euclidean Distance



Clustering using DTW
More Intuitive Clustering

Preliminary Results of DP with DTW





Discussion

Sandia National Laboratories

Benefits of Clustering Algorithm

Arbitrary Clusters	User-Defined Parameters	Multi-Dimensional Data	Noisy Data

Further Work Needed

Time Complexity	Parameter sensitivity	Pruning Metric	Real-Time Algorithm

Conclusion

Sandia National Laboratories

- A pruning algorithm could help drastically reduce the computation time needed to run a large severe accident simulation.
- The clustering of time series can also be used as a tool for regression analysis.
- Future work is needed to refine the algorithm and implement it into real-time tree growth.

Personal: Nevin Martin

Sandia National Laboratories

- B.S. Finance, University of Arizona
- M.S. Statistics, University of New Mexico (Dec. 2016)
- Student intern since August 2015: Org. 6231 & 436
- Hope to continue work at Sandia after graduation.

References

- Diego Mandelli, Alper Yilmaz. "Scenario Analysis and PRA: Overview and Lessons Learned".
- Diego Mandelli. "Scenario Clustering and Dynamic Probabilistic Risk Assessment".
- Diego Mandelli, Peer-Timo Bremer. "Analyzing Dynamic Probabilistic Risk Assessment Data through Topology-Based Clustering".
- D. Mercurio, L. Podofillini, E. Zio, V.N. Dang. "Identification and Classification of Dynamic Event Tree Scenarios via Possibilistic Clustering: Application to a Steam Generator Tube Rupture Event".
- Periklis Andritsos. "Data Clustering Techniques".
- Alex Rodriquez, Alessandro Laio. "Clustering by fast search and find of density peaks".
- Eamonn Keogh, Michael Pazzani. "Scaling up Dynamic Time Warping for Datamining Applications".
- Nurjahan Begum, Liudmila Ulanova, Jun Wang, Eamonn Keogh. "Accelerating Dynamic Time Warping Clustering with a Novel Admissible Pruning Strategy".
- Salah Amean. "Data Mining: Concepts and Techniques".
- Suk Ho Lee, Euncheol Choi, Moon Gi Kang. "Illumination Change Adaptive Tracking Based on Color Centroid Shifting."
- Levent Ertoz, Michael Steinbach, Vipin Kumar. "Finding Clusters of Different Sizes, Shapes, and Densities in Noisy, High Dimensional Data".
- Lance Parsons, Ehtesham Haque, Huan Liu. "Evaluating Subspace Clustering Algorithms".
- Joshua Tenenbaum, Vin de Silva, John C. Langford. "A Global Geometric Framework for Nonlinear Dimensionality Reduction".
- Inyong Seo, Boknam Ha, Sungwoo Lee. "Detection of Sensor Degradation Using K-means Clustering and Support Vector Regression in Nuclear Power Plant".
- Wei Chen, Wei Want, Sandra Batista. "Data Clustering: Algorithms and Applications".
- Pang-Ning Tan, Michael Steinbach, Vipin Kumar. "Introduction to Data Mining".
- Douglas Osborn, Tunc Aldemir, Richard Denning, Diego Mandelli. "Seamless Level 2/ Level 3 Dynamic PRA Clustering".