

Clustering and Visualization of Temperature Time Series

2006 Spring CS235 Project
Danhua Guo
Jianfeng Yang
University of California, Riverside
{jyang, dguo}@cs.ucr.edu

Road Map

- Motivation
- Definition of Problem
- Experiment
- Result and Conclusion

Motivation

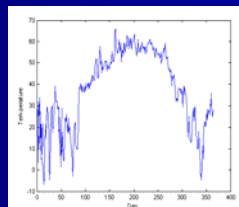
- Try to find an easy way to demonstrate the difference between temperature time series.
- Group 157 cities by temperature, so that one city will lead to another in the same cluster.

Road Map

- Motivation
- Definition of Problem
- Experiment
- Result and Conclusion

Definition of Problem

- Data set: Highest daytime temperature in 157 US cities, from 1995.1 to 2002.12
- Things to do
 - Cluster the temperature time series of each year in each city
 - Visualize the cluster results
 - Make it easy and effective



Road Map

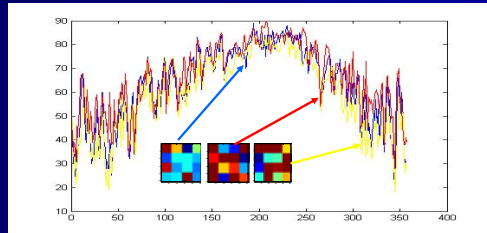
- Motivation
- Definition of Problem
- Experiment
- Result and Conclusion

Road Map

- Motivation
- Definition of Problem
- Experiment
- Result and Conclusion

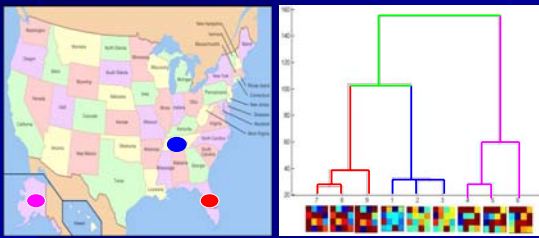
Result

- People cannot always tell the difference between time series
- Icon, in this case, is of great help.



Result (cont.)

- Result overview with geographic verification
- Icon is also shown for each time series



Conclusion

- Hierarchical clustering works well in temperature time series clustering.
- SAX is a good representation of temperature time series.
- Icon and dendrogram show the similarity of different temperature time series in a more straightforward way.

Thanks to

- Prof. Keogh for the idea of SAX and Icon
- Everyone for giving up their sleep in such an early morning

Questions or comments?