

Processing streaming data on novel computing platforms (how hard can it be?)

Eamonn Keogh

Univerisity of California, Riverside

1. ABSTRACT

One of the hottest research topics in recent years is processing streaming data (time series) on novel computing platforms. Such research has applications in wireless health care, robotics, gaming, the automotive industry, and gesture recognition to name but a few. In most cases the platforms in question impose strict constraints on computational resources, producing significant additional difficulties in the already challenging area of time series processing.

In this talk I will argue the following claims.

Similarity search is the fundamental operation for processing series data, and virtually any task, classification, clustering, rule finding, anomaly detection etc., can be efficiently and effectively solved once the similarity search problem is solved.

While there are dozens of alternative distance measures for similarity search, a 50-year old idea, Dynamic Time Warping (DTW) is exceptionally hard to beat.

DTWs often touted lethargy is no more. With four simple new ideas, we can process DTW in real time, even on very low powered devices.

I will illustrate this talk with some experiments on datasets from a diverse variety of real word problems.

2. SPEAKER BIOGRAPHY



Dr. Keogh is a professor of computer science at the University of California Riverside. His research interests include data mining, machine learning and information retrieval. He is one of the most prolific and most cited authors in the area of time series data mining, where his data representations/algorithms have been used in more than 1,000 research efforts. He is a winner of a Vodafone Wireless Innovation Award, a Bill and Melinda Gates Foundation Grand Challenges Award, and best paper awards at SIGMOD, SIGKDD, ICDM, SDM and other venues.