KLINGLER
College of Arts & Sciences
MARQUETTE UNIVERSITY

Department of Mathematics, Statistics and Computer Science

SPECIAL COLLOQUIUM ANNOUNCEMENT

Higher Dimension Topological data analysis: Exploration into Flexible Applications to Real Scenarios

Hiroyuki Sato Ph.D.

Information Technology Center
University of Tokyo

1:00 p.m., Friday, October 20, 2017

Abstract

As data mining grows its significance in real world information extraction, we need a robust methodology without using rich assumptions. In anomaly detection, statistical methods are very successful when we assume some distribution. However, even when we encounter data with unknown distribution, we must first assume some distribution, which may cause incorrect inference.

Topological Data Analysis (TDA) is a new branch of machine learning based on geometric topology. Given a dataset, TDA tries to find its ``shape," or how the points in data are connected, and how they form clusters. Furthermore, it finds the holes of the shape, which imply some data does not exist, which also characterizes the shape.

In this talk, we apply TDA methodologies to anomaly detection. We also apply Morse theory together with persistent homology to the analysis of dynamical system of big data set such as the surveillance video data and the sales dynamics. The idea is that data tendency changes can be observed as a change of data connectivity.

We explore the theory of critical points by using Morse theory. Growth of each dimension shows that some irregular event may have happened on the corresponding date point. Although we need refinement on the interpretation of the result, anyway we can collect the candidates of irregularities. In the talk, we also give some justification of the analysis by using a conventional statistical method. Furthermore, we discuss the limitation or some challenges of TDA.

In TDA, low-dimension homology has been focused on, which is easy to compute, but the applications are limited in nature. Morse theory essentially uses high-dimension homology groups. Their computation is still challenging. Accuracy in some sampling methods is also analyzed in the talk.

1313 W. Wisconsin Avenue, Cudahy Hall, Room 401, Milwaukee, WI 53201-1881.

For further information: http://www.mscs.mu.edu/mscs/resources-colloquium.shtml
or contact Dr. Daniel Rowe #414-288-5228, Daniel.rowe@marquette.edu.

POST-COLLOQUIUM REFRESHMENTS SERVED IN ROOM 342 AT 2:00 P.M.