



<b>TOPIC</b>	<b>A Novel Clustering Algorithm Based on Fitness Proportionate Sharing</b>
<b>ORGANIZERS</b>	Student Leadership Council and Faculty of ACIT Institute and TECHLAV Center
<b>AREA</b>	Machine Learning, Data Analytics, Testing and Evaluation of LSASV
<b>SPEAKER</b>	Xuyang Yan
<b>DATE</b>	Friday July 14, 2017
<b>TIME</b>	3:30 – 4:00 P.M. (EST)
<b>VENUE</b>	Fort IRC 410, North Carolina A&T State University
<b>FEES</b>	No Charge

## SYNOPSIS

Data partitioning is a useful tool, widely used in the analysis of data streams. Data streams refer to a large volume of continuous data that arrives at a high speed and is widely used for the information processing and decision making in robotics and cybersecurity. Through partitioning, data streams can be decomposed into several smaller subsets and a simpler local model will be established for each subset, which makes it easier to summarize the characteristic of the data. Also, the classification can be performed based on the local models, which can be used to develop the speed-up of the learning and decision making processes.

Among all data partitioning techniques, clustering is most widely used. As an unsupervised learning technique, clustering partitions the whole data space into different subspace based on their similarities, which makes it easier to capture the characteristics of the system and data streams. Existing clustering techniques primarily rely on prior knowledge about the data, such as the number of clusters and radii. However, in real applications, the number of clusters and the radii of clusters are usually unknown. Therefore, the performance of clustering methods with overlapping data is degraded due to their limitations in finding all cluster centers with uneven density values.

We have proposed a new clustering algorithm based on the fitness proportionate sharing that maps this clustering problem into a multimodal optimization problem. This new clustering algorithm can effectively handle data streaming without requiring any prior knowledge and provide high accuracy. To avoid the appearance of the redundant clusters, a procedure of niche expansion is employed for the merging of clusters in this novel clustering algorithm. Simulation results for both synthesis data sets and real data sets are presented in this study and reveal that the proposed clustering algorithm based on the fitness proportionate sharing can provide a high accuracy performance without any prior knowledge of data.

## ABOUT THE SPEAKER



Xuyang Yan received his bachelor degree in Electrical Engineering both from North Carolina A&T State University and Henan Polytechnic University in 2016. He is pursuing his master degree at North Carolina A&T State University, and a current student member of the Testing, Evaluation and Control of Heterogeneous Large-scale systems of Autonomous Vehicle (TECHLAV) center. His research area is to apply data-driven techniques in the testing and evaluation of the emergent behaviors of LSASV.