Multi-Scale Graph KRS

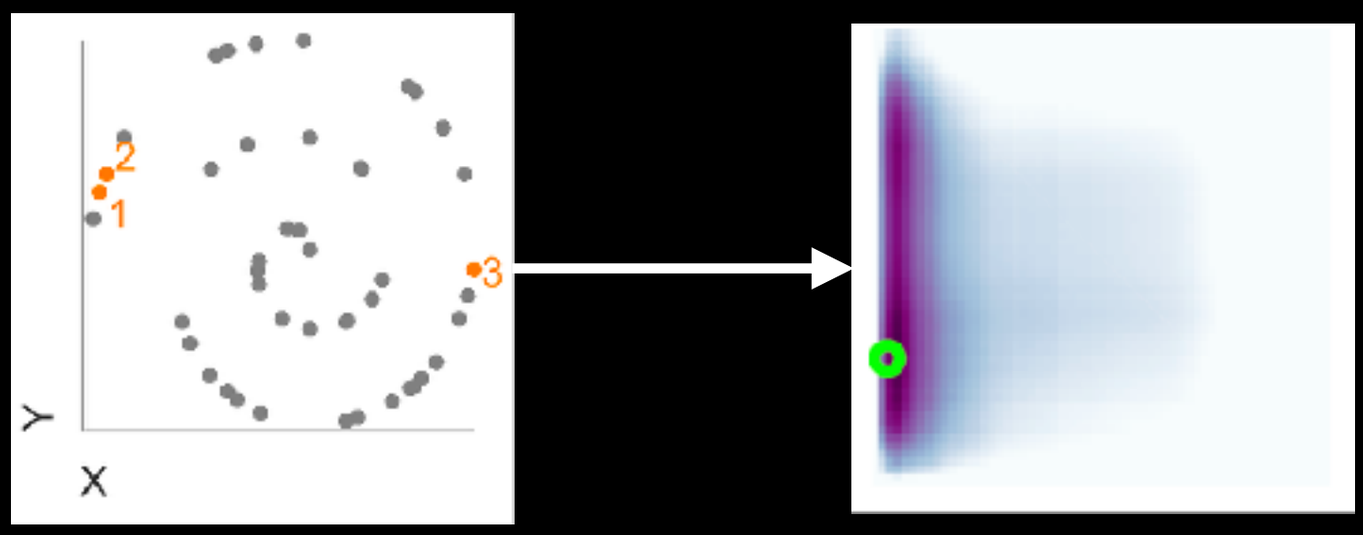
Basic Idea: Many approaches use an graph-based knowledge representation that requires a choice of scale. We *efficiently check all scales and find the best one*.

Breakthrough:

1. achieve the same performance levels as previous methods with provably *fewer samples*,
2. quantitatively characterize the *multiscale geometry* of the function (regardless of types of data or relationships)

So Far: We have *proven and demonstrated empirically superior performance* to current state of the art, including independence testing, equality testing, clustering, topology vs. attribute testing, and feature selection.

DoD Success: Mitigate *batch* effect in fusing multiple disparate heterogeneous data to improve predictive accuracy on a DoD relevant dataset.



V-22 Dataset

Hard Stuff

- We have not yet extended our framework to support multivariate time-series data
- If each event is “categorically” different, pooling across epochs will be difficult
- Some data over-written due to short memory buffers
- Not obvious there exists any calibration data

Success looks like

- We can predict which parts require maintenance in advance.
- A lesser success is we can tell you that *something* is not working, and therefore the plane should get a maintenance check.

Social Network Dataset

Hard Stuff

- It may be a set of networks/ actors with partial or missing alignment across games
- Finding an appropriate notion of “similarity” in the data

Success looks like

- We can find “caricatures”of individuals that most participants are like
- We can assign individuals into groups that act “similarly”

A Few Additional Proposed Datasets

- Something from D3M: as we have no background doing time-series or NLP with MGC, any of the proposed challenge problems are a reasonably large pivot.

- Given spatial transcriptomics data from multiple different labs using disparate experimental modalities, find a “concensus clustering” fusing all of the missing & noisy data with batch effects. This is the 1st challenge associated with the International BRAIN Initiative.

- Given psychological evaluations of 1000’s of individuals including various questionnaires, physical exams, and more, acquired from different labs/hospitals, find “types” of patients that respond effectively to different treatment plans. This is the biggest grand challenge in modern clinical psychiatry.