VALÉRIE POULIN, Tutte Institute for Mathematics and Computing (CSE)
Hypergraph exploration via vectorization

Representing data in a relational manner has become a very common decision for analyzing data sets. In some situations, hypergraph structures are preferred over graphs as they allow for capturing more complex relationships amongst objects. In this work, we present a joint vertex and hyperedge vectorization strategy. The joint vectorization we propose is performed in two steps. We first create vertex vectors based on co-occurrences of vertices in hyperedges. Once vertices are embedded in a space, we consider each hyperedge as a distribution over the vertex space and define a hyperedge distance using a distribution metric. The distance between hyperedges is therefore not limited to the intersection size: two non-overlapping hyperedges can end up being similar if the vertices they contain are.

Through a concrete example we demonstrate how this vectorization allows for visual exploration, cluster interpretation and much more. This analytical framework perfectly illustrates what guides unsupervised data science research at the Institute.