INTRODUCTION

Market basket analysis is the science of discovering customer purchase behavior in order to design marketing and operations management strategies.

There are dozens of enterprise solutions to market basket analysis, based on FP-Growth algorithms. However, they lack robustness and are computationally expensive.

Our proposed solution is based on the graph of transactions, which updates dynamically in time. A real-time controller monitors critical measures of the system and reports any change in customer behavior.

The output can be fed directly to other subsystems for product assortment planning, inventory planning, pricing, promotions, etc. to take suitable actions.

OBJECTIVES

Develop robust algorithms to identify frequent-item-sets.

This algorithms must be:
- Dynamic (handle temporal patterns)
- Fast (computationally efficient)
- Work with real time data
- Handle large scale data

Secondary goal is to develop controllers for real-time monitoring of the vitals of the retail system as they evolve in time.

The controller must:
- Detect changes (e.g., purchasing patterns) as soon as possible
- Be trained to predict changes
- Separate signal from noise
- Recommend actions to take

MATERIALS & METHODS

A summary of the steps of the method is shown below.

- Generating the graph of items
- Community detection using Louvain
- Real-time measurement of the vitals of the network
- Reporting vitals to the controller
- Mining vitals to extract abnormalities
- Deciding on set of actions
- Feedback to the system and repeating the process

Initializing the graph of items is the most important and innovative part of this research. Graph is formed by following these steps:

The community detection, which is the most time taking step does not need to be performed at every heartbeat. Only when the controller decides that current communities are no longer accurate, it sends a recalculation command to community detection kernel.

RESULTS

Measurements are quick and easy. Their complexities grow linearly in problem size. Data mining stage compares newly arrived transactional data with historical data in to match the trends with its knowledge base. Based on the results of matching several set of actions are possible.

To test controller performance, new patterns were injected into streaming datasets. Controller could detect shifts within few iterations. Below is a detection example when a particular cluster has lost an SKU.

CONCLUSIONS

- An effective graph based automated algorithm to find communities of frequently bought together SKUs is introduced.
- Algorithm uses parallel and cluster computing capabilities to reduce execution time. Exhibits good performance on synthetic datasets.

Next Steps for MBA:
- Adapt method to monitor substitutable and complementary products and sales events
- Further operationalize method and enhance robustness
- Investigate Customer-Product Graphs

SPONSOR

CONTACTS

Elham Nosrat, Research Assistant
Cell: 333.353.2159
eno7573@wayne.edu

Ratna Babu Chinnam, Ph.D., Professor
Director - Big Data & Business Analytics Group
Tel: 313.577.4846
Ratna.Chinnam@wayne.edu