Thesis Proposal Documentation

Buctolan, Marjorie
Teves, Shannon Cates

January 20, 2017

Contents

Abstract

Recommendation System has been popular nowadays it has been used for well-known e-commerce sites and software applications like Amazon, Lazada, Netflix, and Youtube. It is used to predict the rating or preference that a user would give to an item. The focus of this study is to implement a recommendation system in AgriSenso, an agricultural web-based application that help and connect buyers/consumers to the sellers/farmers and vice versa. It is also a guide for managing a farm and for growing farm products. The researchers identify different approaches of algorithm techniques used by recommender system such as Collaborative Filtering we used Item-Based Collaborative Filtering Recommendation Algorithms to compare the items that the two target users had rated and compute the similarities of the items ranking. A pathfinding algorithm, A* Search Algorithm was also used for searching all possible nearest seller/farmer from the buyer.

Chapter 1

Research Description

Recommender systems are software applications that make product/service recommendations with the aim of optimizing some user-oriented objective in light of inherent uncertainty over users and contents (Shengbo Guo). Recommendation is widely used technique to guide user to choose the right product online as it becomes the essential feature for the successful E-commerce application. The modern E-commerce system widely used good recommendation system which will not only guide user to choose correct product at correct time but also it helps attracting the customer to do business online effectively. Recommender systems increase interaction to provide a richer experience of the users. To create recommender systems, a reliable algorithm is needed to implement.

1.1 Background of the Study

The AgriSenso system is a web-based application that helps seller to find potential buyer of their products regularly. It is a buy-and-sell platform that connects both seller and buyer through the system or simply via chat or phone call. Also, a guide for managing a farm and for growing farm products. Moreover, a recommender system will be implemented in AgriSenso system. Collaborative filtering (CF) at a recommendation is based on a model of prior user behavior. Item-Based collaborative filtering algorithm is an approach of CF that will recommend the best seller, best buyer, and best product. This approach looks into the set of items and then selects the most similar items. A* algorithm is used to search paths where it moves from the starting point to the goal to find the length of the shortest path. This method will recommend the nearest location of the seller/farmer from the buyer.

1.2 Statement of the Problem

As a result, in the conducted survey to know the current state of the farmers in Iligan City specifically in rural areas and also to the potential buyers of the area, researchers found out that the most common problems that they have faced are as follows:

Seller/Farmer:

a. Farm product underpricing offered by buyers.
b. Difficulties on finding potential buyers of their farm products.
c. Difficulties on selling farm products.
d. Doesn't have a guide to manage their farm products and they don't know what are the best farm products they will raise at a specific time.

Buyer:

a. Difficulties on finding suppliers/sellers to provide them farm products that they want.
b. Difficulties on finding the nearest suppliers/sellers.
c. Difficulties on finding other products that may be useful to the buyer.

1.3 Research Objectives

1.3.1 General Objectives

To implement different approaches of algorithm techniques used in AgriSenso's recommendation system. A* Search algorithm technique finds the nearest location of the buyer to the seller. Item-Based collaborative filtering technique recommend product to the user through the similarities of the purchased item of the other user. Individual user's rank to each item will be used as a data to calculate the similarities of each item that each user purchased.

1.3.2 Specific Objectives

The following are the specific objectives of the research:

1. To conduct a survey to local farmers especially in rural areas to gather informations.
2. To build a web-based system 'AgriSenso' the features to be implemented will only focus on those necessary features affected by the recommender system.
3. To implement the best seller recommender system, where buyers can be recommended sellers based on the biggest number of farm products bought from its different buyer wherein CF algorithm in Item-based approach can be applied.
4. To implement the best buyer recommender system, where sellers can be recommended buyers that has the highest rating given by the sellers wherein CF algorithm in Item-based approach can be applied.
5. To implement the best similar product recommender system, where buyers can be recommended products that has similar preferences of items based on the rating patterns of items given by the buyers wherein CF algorithm in Item-based approach can be applied.
6. To implement the nearest location recommender system, where seller's nearest location can be recommended by buyers and vice versa wherein A* search algorithm can be applied.

1.4 Scope and Limitations of the Research

Recommender system(RS) will be implemented in this research. These are recommendations for the best product, best seller, the best buyer, and the nearest location of the seller from a buyer. The recommender system for the best seller will be based by the number of farm products bought by the buyer. The researchers will get data from its system database. Therecommender system for the best buyer and best product will be based by the rating given by the seller. To control the data gathering process, the researchers will create an account for each test user(seller) and will provide the list of buyers and list of products that they will going to rate. An interview will be conducted to each test user so that they will have an idea as to what specific data to provide. The test user will rate a buyer from one star to five stars, one star will be the lowest and five stars will be the highest. The recommender system

for the nearest location will only be provided locally in the Philippines. Also, test user will be needed in this RS. Test user should provide his/her exact location.

1.5 Significance of the Research

AgriSenso system serves as a bridge for both sellers/farmers and buyers. Basically, it fills the gap between farmers and agriculture experts. This research generates recommender system that can provide user a scalable way to help discover what product they may like and who is the good or maybe the best seller or best buyer to get accounted with. Therefore, this research can help users make good and faster decisions.

Chapter 2

Review of Related Literature

2.1 Item-Based Collaborative Filtering

Due to the tremendous growth in the amount of available information and the number of visitors on Websites, new recommender system technologies are needed that can quickly produce high quality recommendations, even for very large-scale problems. To address these issues Sarwar, Karypis, Konstan, and Riedl explored item-based collaborative filtering techniques. The research paper focused on the analyzation of different item-based recommendation generation algorithms and different techniques for computing item-item similarities. The researchers evaluate a new algorithm for CF-based recommender systems and compare the results to the basic k-nearest neighbor approach. According to the researchers their experiment suggest that item-based algorithms provide better performance and quality than user-based algorithms for item-based techniques hold the promise of allowing CF-based algorithms to scale to large data sets. Another study, Item-Based Top-N Recommendation Algorithms of Mukund Deshpande and George Karypis. The emergence of e-commerce sites and rapid growth of the World Wide Web has led to the development of recommender systems. User-based collaborative filtering is said to be the most successful technology for building recommender systems and is used in many commercial recommender systems. Unfortunately, the computational complexity of these methods grows linearly with the number of customers in commercial applications whichtypically runs several millions. Their article presented one such class of model-based top-N recommendation algorithms that use item-to-item similarities to compute the recommendations. The result of their experimental evaluation on eight real datasets shows that "item-based algorithms are up to two orders of magnitude faster than the traditional user-neighborhood based recommender systems".

2.2 A-Star(A*) Search Algorithm

In the study of Xiao Cui and Hao Shi, A*-based Pathfinding in Modern Computer Games explores the relationship between various A*-based algorithms and reviews a number of popular A*-based algorithms and techniques according to the optimization of A*. They mention some ways to improve the performance of A* search include optimizing the underlying search space, reducing the memory usage, improving heuristic functions and introducing new data structures. The researchers also define another way to make some contribution to the game AI community to the real computer games. The research paper of Hart, Nilsson, and Raphael they described how heuristic information from the problem domain can be incorporated into a formal mathematical theory of graph searching. And they demonstrated an optimality property of a class of search strategies. The study concerned with the subgraph G from some single specified start node s. An algorithms that search G to find an optimal path from s to a preferred goal node of s.

2.3 Recommender Systems

According to Beel, Gipp, Langer, and Breitinger more than half of the recommendation approaches applied content-based filtering (55the reviewed approaches, and Graph-based recommendations by 16revealed some shortcomings of the current research such as it remains unclear which recommendation concepts and approaches are the most promising. Different results on the performance of content-based and collaborative filtering. They also discussed three potential reasons for the ambiguity of the results. First is the several evaluations had limitations based on strongly pruned datasets. Second, some authors provided little information about their algorithms, resulting for the difficulties of re-implementing the approaches. And lastly, minor variations in datasets, algorithms, or user populations inevitably lead to strong variations in the performance of the approaches.

Chapter 3

Research Methodology

This chapter discusses the relevant theories and concepts used in this research.

3.1 Survey

3.2 System Architecture

3.3 Item-Based Collaborative Filtering Algorithm

3.3.1 Cosine-Based Similarity

The similarity between the two items are measured by computing the cosine of the angle between the two vectors (where two items are two vectors in the m dimensional user-space).

3.3.2 Correlation-Based Similarity

The similarity between two items i and j is measured by computing the Pearson-r correlation.

3.3.3 Adjusted Cosine Similarity

Offsets drawback by subtracting the corresponding user average from each co-rated pair.

3.4 A-Star(A*) Search Algorithm