SayMyWave: Website Recommendation with Collaborative Filtering

Eren Golge

Tunc Gultekin

Ilker Sarac

Ahmet Iscen

Outline

- Goal
- Motivation
- Related Work
- Methodology
- Our System
  - Recommendation Engine
  - Database
  - Web Page
  - Browser Extension
- Prototype Demo
- Conclusion

Goal

▸ Recommending new web contents and services to users, based on their previous web history, and activity.

Motivation

- There are billions of web pages today.

- The users have many options to choose from when they need to access an information.

- It is important to guide the users to websites that they are likely to enjoy

Motivation

- Knowing user preferences are important for targeted advertisements.

- Such recommendation system can provide a baseline for targeted advertisements.

Related Work

▸ Recommendation systems have been examined for a long time;

◦ Tapestry, Video Recommender and Ringo
  • Uses closely related groups to understand similarity of the instance

◦ Bayesian Network Based Systems
  • Not flexible, requires updates when user preferences changed

◦ Clustering Based Systems
  • Requires too much time for cluster construction.

Methodology

- Web content requests are being captured and sent to recommendation engine.

- Similar users are found in the recommendation engine.

- Web content url, that are visited by similar users, is sent to clients.

Our System

▸ Overall our system has four main components;

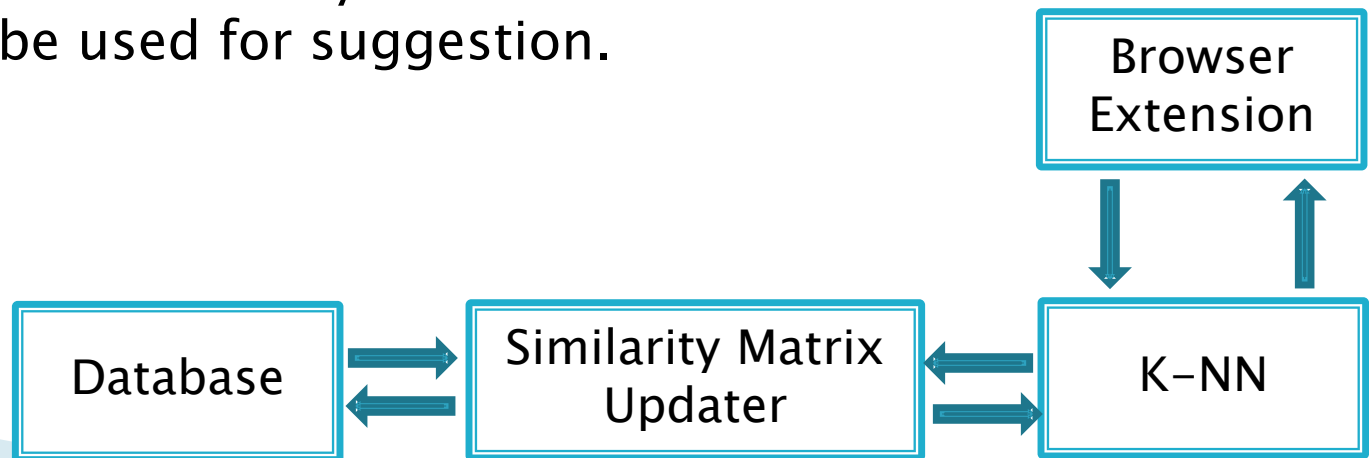| Browser Extension	⇄	Recommendation Engine	←	Web Page

Database
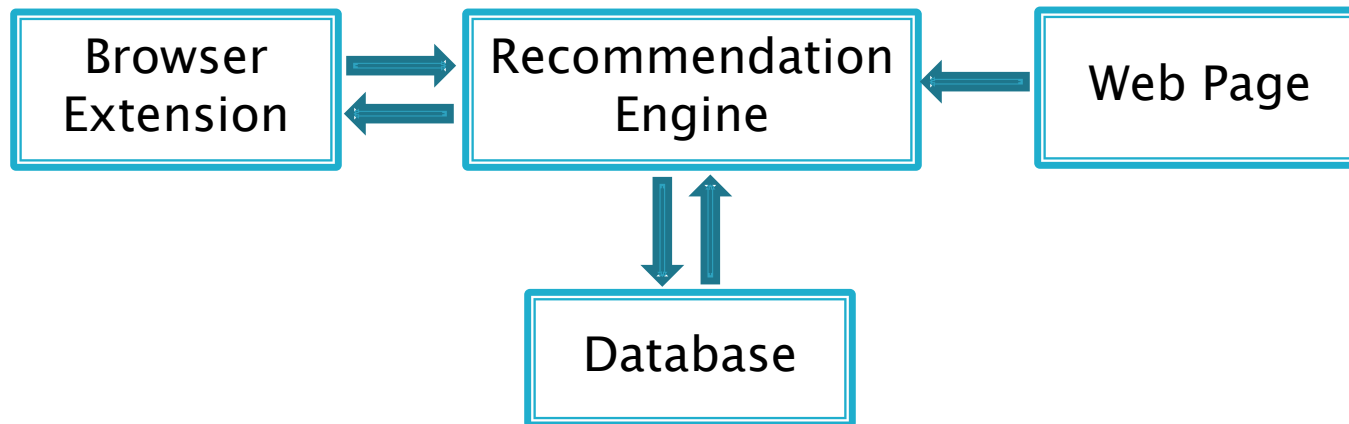
Recommendation Engine

- The core of SayMyWave is the recommendation engine that processes all the data in order to provide suggestions for specific user.

- An accurately functioning engine provides better circulation of user.

- Larger crowd is meant to be broader wisdom.

Recommendation Engine

▶ General progress of the engine is as follows;

- ◦ Collect data from database in the form of a matrix.

- ◦ Fill the missing values of each user by finding the most similar two users.

- ◦ Apply Cosine Similarity measure to the data and cache the result to be used for suggestion.

```
┌─────────────┐
│   Browser   │
│  Extension  │
└─────────────┘
      ↕
┌──────────┐   ┌──────────────────┐   ┌────────┐
│ Database │ ⇄ │ Similarity Matrix│ ⇄ │  K-NN  │
│          │   │     Updater      │   │        │
└──────────┘   └──────────────────┘   └────────┘
```

Database

- Database will constantly interact with the Extension and Recommender system.

- It will keep the entire user and website information.

| user_id	site	visit_c	last_vis	typed	hidden	frec

Web Page

- Web page of system will be provide following features;
    - Registering new users.
    - Suggesting web sites to user without using browser extension.

- We have not implemented it yet.

Browser Extension

▸ 2 different ways were planned and tested for capturing user behaviour.

◦ Http Port Listener Application
  • Captures all of traffic, most of them are irrelevant for user web site requests.

◦ Browser Extension
  • Practical and only captures users' real web requests

Browser Extension

▸ Chrome browser extension was implemented. Behavior of it as follows;

◦ Reads User Id from a cookie.

◦ Captures all the URLs that are entered into Chrome's address bar with their frequencies and sends them to system.

◦ Whenever a user clicks the button of extension, it suggest a new web site by quering over the system.

Architecture

Chrome Extension.

Internet

Client
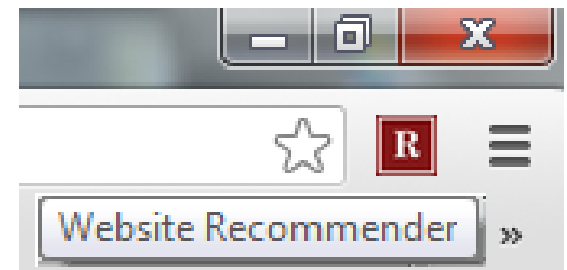
Recommendation Engine
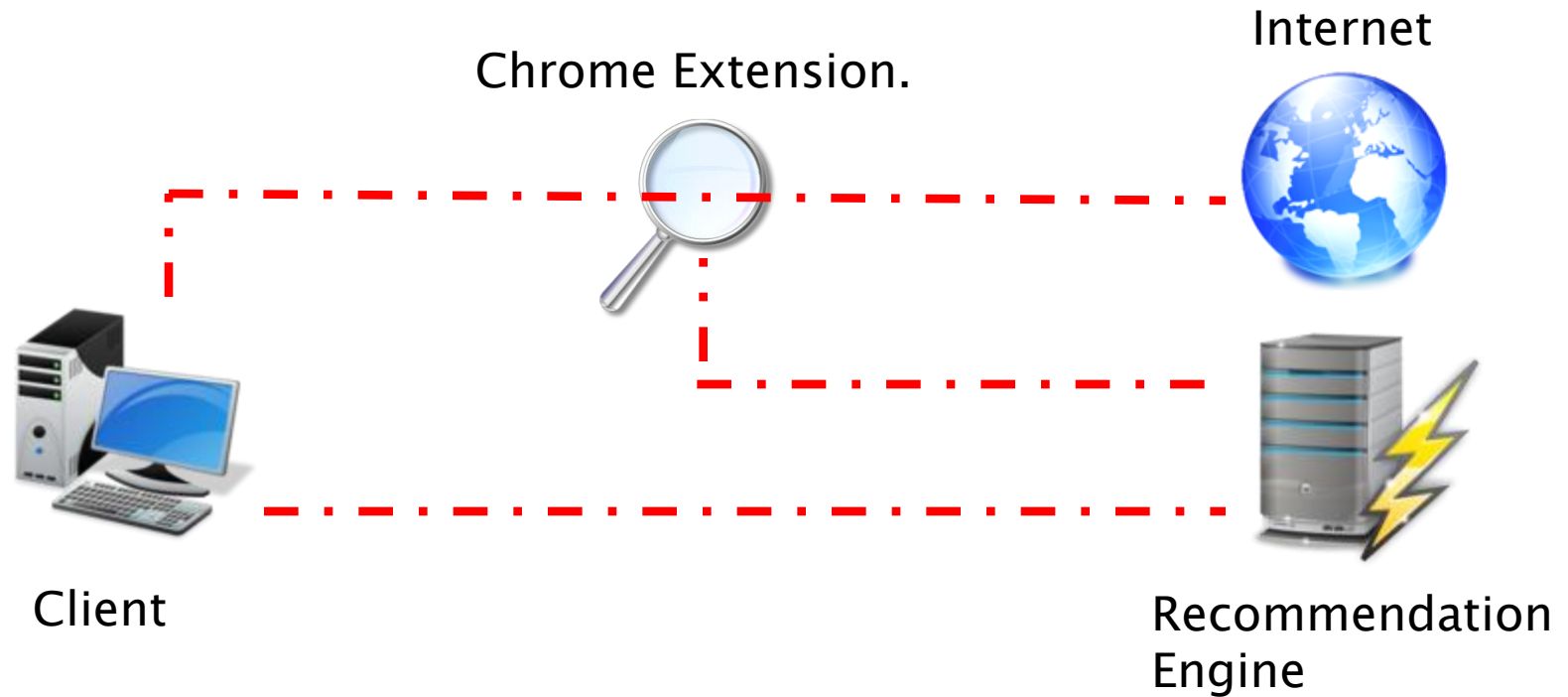
Tools

- Recommendation System
  - Python's Django framework
  - ASP.Net for Prototyping

- Chrome Browser Extension
  - Javascript and Ajax Technology

Prototype

- 10 random users and their web sites were created by using 20K record web history of a person.

- For each incoming query, weighted cosine similarities are calculated.

- Most similar user is found.

- One of the entered sites of similar user is retrieved.

Conclusions

- Our recommendation engine actively is at the development phase.

- Our prototype of the system is designed and browser extension is ready to process data.

- We are expecting to have fully utilized and efficiently working web suggestion engine.

Thank You