‘LUPPEIT’
WEB & MOBILE APPLICATION FOR GATHERING RECOMMENDED USER RELATED DATA

by Faruk Kuşcan

Submitted to the Department of Computer Engineering
in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Engineering

Boğaziçi University
November 2012

TABLE OF CONTENTS

ABSTRACT

Project Name	:	LuppeIt
Project Team	:	Faruk Kuscan
Term	:	2012/13 I.Semester
Keywords	:	Social Media Analysis, Recommendation, RSS
Summary	:	LuppeIt is a tool for any kind of user who wants to find links in internet which he/she would be interested in more easily. Also it is a good social media analysis tool. LuppeIt is a web project which a user can be registered. For registered user, LuppeIt keeps the track of all user actions in the application. According to this data, LuppeIt recommends user related links shared by other users and found from RSS feeds.

1.INTRODUCTION

LuppeIt is my senior project. In this part, the problem and my project (which is the tool to use to solve the problem) will be analyzed.

1.1.The Problem

The internet has lots of news, shares, blog posts and many other different kinds of information in it. However, it is also a very big sea so sometimes we can not find what we really need in the internet. When someone needs to find something in internet, search engines are used like Google. This is a way of searching something in internet when we “need” that information.

Also there is another kind of search. That is when we do not need that specific information. For example, we type ntvmsnbc.com and see what news are posted today. However, there are hundreds of news posted each day. Most of those news are not interesting for all users of ntvmsnbc.com. In this part the other kind of search starts. That is a manual search for that specific user to find content which he or she will be interested in to read.

This is not a good way to find the content what the user is interested in. In the first part, the user should filter all the web by opening a web page. The web sites are not many. If the user doesn’t know abcd.com, he or she will not be able to see the content he or she may be interested in. Also in the second part, the user should search manually by clicking on categories etc.

The problem here has two main points. First one is the following: Users can not know all the web sites in the internet where the content they are interested in are published. Second point is the manual search made by user to find the content they are interested in. Their interests are surrounded by many other content which they do not want to read.

1.2.What is LuppeIt?

LuppeIt is a web application which gathers content from all around the web for its users and filters these content according to its users' interests. LuppeIt does this filtering job by given data from the user and all the actions made by that user on LuppeIt.

As the problem is analyzed in the first part, LuppeIt finds solution to the problem as told in the following.

The problem had two points. First one was that the user does not know all the web pages in the internet. LuppeIt gets content from many web pages. If these web pages have RSS feeds, LuppeIt is integrated to that web page. If it does not have an RSS feed, singular links of that web page can be shared by users of LuppeIt. All users of LuppeIt widens the web pages LuppeIt knows. When LuppeIt knows a web page, its users will know that page if content of that page complements that user's interests.

The other point of the problem was the manual search part. In LuppeIt, when a user is registered, his or her interests are asked to him or her. User selects the categories that he or she would be interested in. After registration these interests may change and the user is able to change his or her interests that is told to LuppeIt. Also, LuppeIt keeps the track of all of its users. If a user likes or dislikes a content, LuppeIt stores this action as a data to use. Even if a user just views a content, LuppeIt stores this action as a data too.

In this way, LuppeIt tries to know its users and its users' interests. The manual search made by users is not an action in LuppeIt. Because, LuppeIt does this filtering automatically. LuppeIt knows what each user would be interested in and exposes only that content to each user.

1.3.LuppeIt's Basics?

Share

Share is either a url submitted by a user from the internet or a url which came from an RSS feed that LuppeIt uses as resource.

Category

In LuppeIt all shares are seen under a category like "World News, Politics, etc." These categories can not be defined by users. A share can not be related to multiple categories.

Tag

Relevant words about a share. A share can have multiple tags. Users can add new tags. Tag has a truth value in LuppeIt for finding better tags for shares.

Resource

The source web page of a share is resource. As an example, if a share's url is http://www.ntvmsnbc.com/id/12345, the resource is "ntvmsnbc.com".

Like/Dislike (Luppe/Dig)

Users will be able to like or dislike a share. This will help LuppeIt to learn more about the user and more about the average of all users and resources.

Interest

Users will be able to mark some categories as their interests.

These are basic definitions that we need to know to understand LuppeIt.

User Types

There are two types of users on LuppeIt: Anonymous user, registered user. For registered users, LuppeIt will look to their user past and behave according to that.

However, for anonymous user, LuppeIt will act according to the general average of all users likes and dislikes.

2.Purpose of LuppeIt

2.1.Purpose

Purpose of LuppeIt is to help its users find relevant news, blog posts and any other links in the internet in a manner that they choose from LuppeIt ‘s features. They can search for a keyword and get a result set according to their past usage in LuppeIt or they can search for a result set with some parameters like “location, time, etc.” While a user uses LuppeIt, LuppeIt learns about that user. In that way, LuppeIt learns about the user more and that user can find what he/she likes more in LuppeIt.

2.2.Social Media Analysis

Social Media Analysis is a very hot topic nowadays. Many companies pay lots of money to make social media analysis. LuppeIt works for its users. Also, LuppeIt works as a tool of social media analysis.

LuppeIt has demographic information of its users. LuppeIt knows how old are its users, or where are they from etc. Their actions in LuppeIt makes sense when these actions are gathered together.

LuppeIt uses these information again with its users. Users are able to do search for a keyword. They are able to do a search with criteria like location, age, etc. When a users makes a search for “sports” in LuppeIt and gives criteria of location Turkey, most viewed, most lupped, most digged shares will be gathered together as a search result set to the user. LuppeIt has this kind of a search mechanism tool to make very fast and useful social media analysis.

3.REQURIEMENTS

3.1.Functional Requirements

LOGIN / REGISTRATION

Users will be able to login to the system. The whole system will work according to user data when a user is logged in. If the user is not logged in an anonymous user will be using the system. However, if the user wants to find his interests from the web, he or she must register and log in.

SUBMIT LINK

A user will be able to share a url with choosing a category and giving some tags to that url. A user must be registered and logged in to submit a link.

SEARCH FOR A KEYWORD

A user will be able to search for a keyword. He or she will be able to add criteria to that search. A user will be able to add location and time criteria to a search. Result set will be user related content if no criteria are given. If criteria exist, result set will be gathered according to a hypothetical user which those criteria represents.

VIEW / LUPPE / DIG A CATEGORY / SHARE

A user will be able to view a category. In that category page, user relevant content of that category will be gathered. The user will be able to click on a share in that category and view that link. Also, the user will be luppe (like) or dig (dislike) those shares. View, luppe and dig actions will be stored by LuppeIt to be used to know the user better. Result sets created for search, category contents for user etc. will be generated according to these data.

CHANGE INTERESTS

A user will be able to tell his or her interests during registration. Also after the registration he or she will be able to change his or her interests.

3.2.Non-functional Requirements

- Availability
    - The application will be available on web interface.
    - The application should run 24/7.
    - Users from all around the world must be able to use it in English.
- Performance
    - Recommendations must be gathered and exposed to each user in a short time.
    - Cache levels must be used to fulfill this requirement.
- Security
    - Users' login data must be stored in a secure way.
    - Users' demographic data must be stored in a secure way.
    - Tag mechanism must be implemented in a way that users can add tags to shares etc. However, tags must be securely controlled to avoid misleading tags.
    - Users must be able to share links from internet but advertisements must be forbidden securely.
- Usability
    - LuppeIt must have an easy to use user interface.
    - Luppe and Dig actions must be analyzed in a page to the users.
- Technical
    - Play! Framework will be used in implementation. (An MVC framework works with Java)
    - MySQL database will be used as data storage.

4.HOW LUPPEIT WORKS?

4.1.Working Environment

LuppeIt is being implemented with Java. An MVC framework called Play! is being used. This MVC framework is connected to a MySQL database. The application is deployed on Play! Framework's built-in application server. The current version of LuppeIt is deployed on a live server.

http://luppeit.com

http://swe.cmpe.boun.edu.tr:7880

4.2.Algorithms for Search

Search without criteria:

Take the query input;

Take the user history;

Take the user interests;

Merge the user history and interests with the query;

Create the result set

Search with criteria:

Take the query input;

Take criteria;

Create a model user with criteria;

Create the model user's interests;

Merge the model user's interests with the query;

Create the result set

4.3.Algorithms for Recommendation

Anonymous user category page result set:

Create a model user with all users' average;

Take the category;

Merge the category with model users' interests;

Create the result set;

Registered and logged in user category page result set:

Take the user history;

Take the user interests;

Take the category;

Merge the category with user's interests and history;

Create the result set;

5.CURRENT PROGRESS

LuppeIt is deployed on http://luppeit.com . Current progress of the project can be viewed from the url. Also the code is on github and progress can be viewed from there too. https://github.com/kuscan/luppeit

6.CONCLUSION

LuppeIt is a good tool to find content that you are interested in the web. It is being implemented right now. When it is done, it will be an easy to use web application. According to the implementation progress, if I have time, it will be implemented on iOS and Android mobile platforms too.