Web Mining Final Report

Team 14: Watch Your Back

05/13/2014

Ching-Tang Lin
Su Zhao
Yuan-Jen Lee
Yufeng Mao

WEBSITE ADDRESS

Contents

Introduction .......................................................................................................................... 3

System Objective ................................................................................................................ 4

Competitors and Market Research ................................................................................. 4

System Design .................................................................................................................... 5

Unique Features/Novelty .................................................................................................... 12

Revenue Model .................................................................................................................. 13

Future Directions .............................................................................................................. 13

Role & Contribution ......................................................................................................... 14

References ......................................................................................................................... 14
Introduction

The reason we came up with sex offender idea was because there were several sexual crimes happened in January 2014 in China. Children were the most innocent victims in those cases. Chinese society outraged by these sexual crime cases, so media began to discuss how to protect children and minimize sex offenders’ potential risk to society again.

Based on the study from Bureau of Justice Statistics, compared to non-sex offenders released from State prisons, released sex offenders were 4 times more like to be rearrested for a sex crime (Durose, Langan, Schmitt, 2003). United States is also facing the risk from sex offenders, so we decided and developed the project to protect our users from sex offenders.

Our website name is Watch your back, which reminds our users to be careful of our surround environment to be away from sex offenders. The website has the functionalities to locate sex offenders, teach users how and what can protect themselves and link to multiple Social Medias to receive up to date sex offender information. Also, with the advent of Web 3.0, our product provides mobile application to achieve SoLoMo for users to better receive information and protect themselves.
**System Objective**

- Build a website to provide easy-accessed, accurate, and intuitionial data to user
- Build a mobile application to achieve SoLoMo features in order to better serve users
- Target college student and people need to relocate
- Keep up to date sex offenders’ data from Department of Justice of each state
- Present sex offender data on the web, such as name, gender, location, etc.
- Achieve Google Map location trucking to find the number of local sex offenders and their detail location
- Effective APIs for user to connect to social media in order to receive most recent sex offender information
- Offer an one stop source for every type of information about sex offender, such as Youtube Video, RSS news from New York Times
- Provide custom Google search option and topics are related to sex offender
- Offer Amazon search option to search self-protection products and provide highly recommended products
- Offer email alert to users who signup their zip code and email address to let them know new moved-in sex offender in their area
- Provide visualized sexual crime data analysis and social media analysis to keep truck of hot topics related to sex offender

**Competitors and Market Research**

Sex offenders’ data are collected by Department of Justice of each state and United States National Sexual Predator Online Website (NSPOW) combines all the data to public. However, state government websites have bad visualization for users to locate sex offenders. NSPOW has delay to receive the up to date data from state governments and it also takes long time to do the ETL process then present it online. The gap brings Watch Your Back a good chance to win in the market with great virtualization and up to date data, which obtained from each state government. Compare to commercial websites, most of them could not provide up to date data and they do not have the functionality to provide web analytics report. For example, Family WatchDog is the big player in the market with good virtualization, but it has too much advertisements on website which destroy the design and customer satisfaction. iCrime Watch provides multi-search options, but the website is too limited functionality which only focuses on sex offender search.
Our team utilizes ten APIs for the website and calls each API through a combination of ASP.NET and Java Script.

IIS7 is used as the application’s web server.

The database that stores all the data is hosted on SQL Server 2012.

The website is housed on a cloud server from Amazon EC2 Web Service, which allows anyone to view our web site.
**Amazon EC2 Web Server:** the website deployed on Amazon EC2 Windows Server 2012 and Microsoft SQL to store sex offender data. It also provides flexibility and easy to use for small business.

**APIs**

- **Facebook (Share & Like):** provides users to like and share their idea from our website. Also, user can like our Facebook homepage.
- **Twitter (Share & Live Tweet Feeds):** provides users to share their idea from our website. Also, user can read live tweet feeds from our sidebar and share their thought from the website.
- **Google+:** provides users to like our Google+ fans homepage.
- **Google Map:** shows users’ current location and sex offenders nearby. Different color dot on screen shows different numbers of sex offenders. Sex offenders’ name and address are shown on Google Map.
- **Google Translate:** provides easy content translate function for users without good English skills.
- **Google Customer Search:** shows results related to sex offenders. It provides more accurate result and save users’ time.
- **YouTube Video Search:** provides user the function to search any videos on YouTube.
- **Amazon Products Search:** provides users personal defense equipment options. If users purchase through mobile application, we earn 6% of purchase. We also recommend some products to users. Users can also share or like the products
- **RSS:** provides live RSS news from New York Times Sex Crime News. It gives users national wide news of sex offenders.
- **Tableau:** visualizes web analysis and present on website.
Here is the Amazon Products Search API. Users can purchase personal defense equipment through this API.

Here is the Google Map API. Users can enter their zipcode to search sexual offenders near themselves.
Mobile App Design

We developed the iOS app in Xcode.
Web Analysis

Database Analysis:
- Cleaned the data using Excel, and made the data that are from the different sources/states consistent to do the further analysis
- Visualized the gender and age distribution of the sexual criminals
- Created a heat map to show the density of existing criminals in a certain area.

![Gender and Age Distribution](image1)

Twitter Analysis:
- Collected more than 4000 Twitter feeds that are related to our topics in three days using Python.
- Cleaned the data using RapidMiner: removed duplicated tweets, missing values
- Selected the highest frequency words in these Tweets, and created a word cloud to use the size of the words to indicate the frequency.
- Picked out the most popular and active Twitter accounts that are related to our topic: the size of the words indicates the number of the followers and the color means the number of the Tweets of the account in a period of time.

![Heat Map](image2)
Unique Features/Novelty

iOS Application
Built a native iOS app based on our website to give people better user experience. The iOS app is created to reflect the concept of Web 3.0/SoLoMo. It can detect your real time location using the GPS censor in your smartphone and show the safety situation around you. In the app, you can purchase the related products we suggested and search for other products you like. And you can also watch videos and news that may help you to gain your awareness of sex offenders. For each portion of the app, the social media-sharing feature is everywhere, because the latest iOS system has many kinds of social media integrated on system level. Whenever you share something, you can add your location and select the audience. The iOS app is built based on the website and beyond the website. It provides the same rich information as our website and better accessibility to users.

Visualization Using Tableau
Our team analyzed both sex offender database from state government and Twitter feeds related to sex offender. We provided clear visualization using Tableau, which is a more intuitive for users to understand the hot topic and some other special and interesting results.

Multiple Social Media
Social media are easier to popularize and increase public awareness of Watch Your Back. Our team created multiple social media home pages for users to follow and like, so they can receive most recent updates about sex offender related topics through social media, which is much easier and faster for them to type our web address in browser.
Revenue Model

- Since our target users are university students, we will ask university and local merchandise to sponsor our website.
- The mobile application provides Amazon product widget, and we can acquire 6% of purchases made through the application.
- The mobile application will be sold on Apple App Store for $0.99 cents per download.
- Our service can be packed and sold to other service websites or applications.

Future Directions

- Currently Watch Your Back has sex offenders’ data from Arizona, North Carolina, Florida and Georgia. The next big step will expend the data coverage to entire United States.
- Currently the mobile application is on Apple iOS platform and we will transplant the application to other platform such as Android platform and Windows Phone platform.
- Currently Watch Your Back is ready to be released to public, improve website visibility is the most important work to do on business aspect. We will put our website on a good search ranking on Google and we will host local college activities to popularize our website and mobile application.
Role & Contribution

• Yuan-Jen Lee:
  o Google Maps API
  o EC2
  o Social Media
  o Data Preprocessing
• Ching-Tang Lin:
  o Front-End Development
  o Back-End Development
  o Integrated 10+ APIs
• Yufeng Mao:
  o Developed the iOS App
  o Designed the Website Layout
  o Tableau Data Analysis
• Su Zhao:
  o Analyzed Criminal Data
  o Wrote Content & Report
  o Gathered Live Tweets

References