

HighFly

Mine/Discover Opportunities in Your Professional Network

Software Requirements Specification Document

Version: 0.1

Date: 09/05/2014

Prepared by: Ahsanul Karim

Email: karim.ahsanul@gmail.com

Github: <https://github.com/akarim78>

1 Introduction

Social media and networking sites are being used for career research, job search and and a tool for marketing one's own skills and expertise to future employers. For students and recent graduates, professional networks (like LinkedIn) become the place where they try to connect to a different group of people than other social networks (like Facebook). They try to connect to alumni communities from their own alma mater, influential people in the same field of study and recruiters from the industry. On the other hand, recruiters also look for potential candidates for jobs. Both the user groups use the professional network as a tool to achieve their goals.

In order to use the information, opportunities and people available in one's professional network, we propose a tool **HighFly** which uses techniques of data science and natural language processing to analyze the network and provide insights through visualization. In this document, we describe the features and functionality of **HighFly** along with development plan and all the technical aspects.

For the prototype, we choose LinkedIn as the target social network for the following reasons:

1. LinkedIn is world's largest professional social network with more than 313 million users in more than 200 countries. LinkedIn claims that professionals sign up to join LinkedIn at a rate of more than two new users per second. There are over 39 million students and recent college graduates on LinkedIn and they are considered as the fastest-growing demographic for LinkedIn [1].
2. LinkedIn provides developer API which also make the collection of data for an individual user much more convenient through API calls [2].

2 Overview

This document is organized as follows:

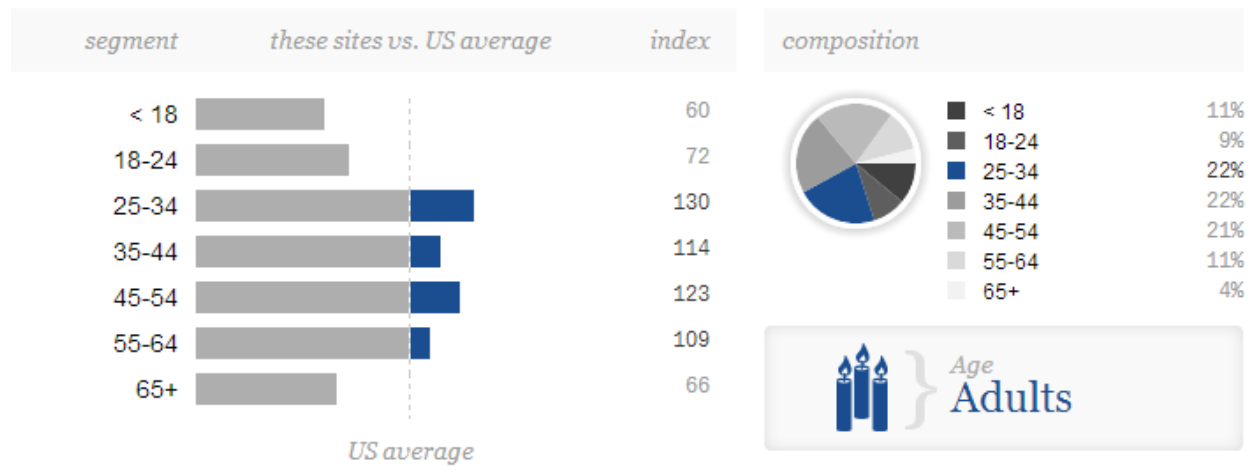
1. Section 3 describes the background and the problem. It also briefly specify what solution **HighFly** is going to provide the users.
2. Section 4 is dedicated to define the application. It shows the list of features along with visual mockups. This section will be open for now and will be updated frequently in the course of the development.
3. Finally, section 5 addresses the technical aspects and choices in terms of technologies. It also provides a development plan for the rest of the semester.

3 Overall Description

This section will present a detailed overview of the background and the problems that we want to solve using the current project.

3.1 Background

In United States, 22% of online adults use LinkedIn to maintain a professional network using LinkedIn [3].



The age group of 25-34 is one of the largest age based demographic that are likely to look for a network which will help them to get a good job. There are also a large number of companies and recruiters look for prospective employees in their LinkedIn network. These networks of individuals bring a lot of data that can be used to help an individual seeking a new job or a recruiter to find potential good candidates. In this project, we focus only to the users that are looking for jobs. We will identify few of the problems that an individual faces during job searching period and try to address them through different features.

3.2 Problem

Recent graduate from colleges and universities create their resume and go through job sites and social networks. They update their profiles to get attention of recruiters and companies. During this process they often face following problems:

1. Creating multiple versions of resume for different job postings according to job requirements and managing them. Often they just keep posting resume without paying much attention to research the job requirements and background of companies due to the desperation. Often they miss to list or focus the key requirements in their resume. Most of the if the resume does not look much relevant to the job requirements, they are often ignored. Managing multiple resume and addressing the job requirements properly is a problem.
2. Another problem is to find out people in the network who can help a user in the process of job searching. Finding a helpful connection (an alumni from same institution or an expert of the field) can often open up new opportunities.

3. Most of the time users miss a lot of job opportunities if they can not search for jobs in different job sites. LinkedIn often sends mail notifications to users, but there are a lot more job sites out there and it is hard to manage so many accounts.

3.3 Proposed Solution

In order to solve the problems we propose a service named **HighFly**. The idea is to help users in following ways:

1. Managing multiple profiles and resumes.
2. Find important connections for their career.
3. Getting job alerts from multiple sources along with relevance score with their profile.
4. Get visual analysis of a job requirement.

3.4 Target Users

The main target user groups right now are the recent graduates from colleges and universities. For first prototype, the target will be satisfy the requirements of a computer science graduate. Features can be generalized for people with various backgrounds then.

3.4 Business Plan

TBD

4 Features

This section provides a list of functionality defined so far. It will be modified and extended according to user feedback.

4.1 List of Functionality

The following features are determined from the high level requirements. Currently the focus is only to the job seeking users. With further development, we can think about a tool for recruiters.

1. Sign up with LinkedIn

User will be able to sign up using LinkedIn account. It will be SSO(Single Sign On).
User will be able to sign in directly after sign up if she is logged in already to LinkedIn.

2. Creating Resume

User will be able to create resume from scratch by adding custom fields. User can also generate resume from LinkedIn profile data. User will be able to keep multiple copy of resume for different job specifications.

3. Importing Profile from LinkedIn Account

User can import all the information from her LinkedIn profile using REST API call and populate her **HighFly** account. This includes only the profile specific data, not the Connection.

4. Importing Connections from LinkedIn

Using REST API calls to LinkedIn user will be able to import all her Connections for profile analysis. Here we will need to explore how many levels of connections can be retrieved.

5. Analyze and Cluster Profiles of Connections

User will be able to see any Connection's profile with analysis. Connections can be clustered based on expertise, endorsements and recommendations. For example if user wants to see only the recruiters for IT industry, she can provide keywords and get recruiters clustered in different types of IT companies (again companies can be clustered based on technologies they want).

6. Visualizing analysis and clusters

One of the main features of the application is to show analysis of the network visually which makes identifying certain types of people easily and more effectively. This feature will involve using different visualization libraries.

7. Analyze User Profile and search Companies and Jobs

According to users multiple version of resume and profile there will be options for searching companies and jobs. Search results will be shown according to relevance score with user's current profile.

8. Analyze job descriptions and customize resume

One of the problems we face during job search is that we often miss important requirements and focus in a job description. Using this feature, user will be able to analyze a job description (fetched from LinkedIn or any other job site) and create a visual representation (may be a tag cloud). This will help the user to understand what should she focus in her resume. Also there can be a scoring mechanism to generate a score of user's profile to a particular job requirement.

9. Use Profile data to fetch jobs from other job sites

There are a number of great job sites which provide developer API's. We will add job stream boxes from different sites (StackOverflow, Dice etc) including freelancing market places (Elance, ODesk etc.).

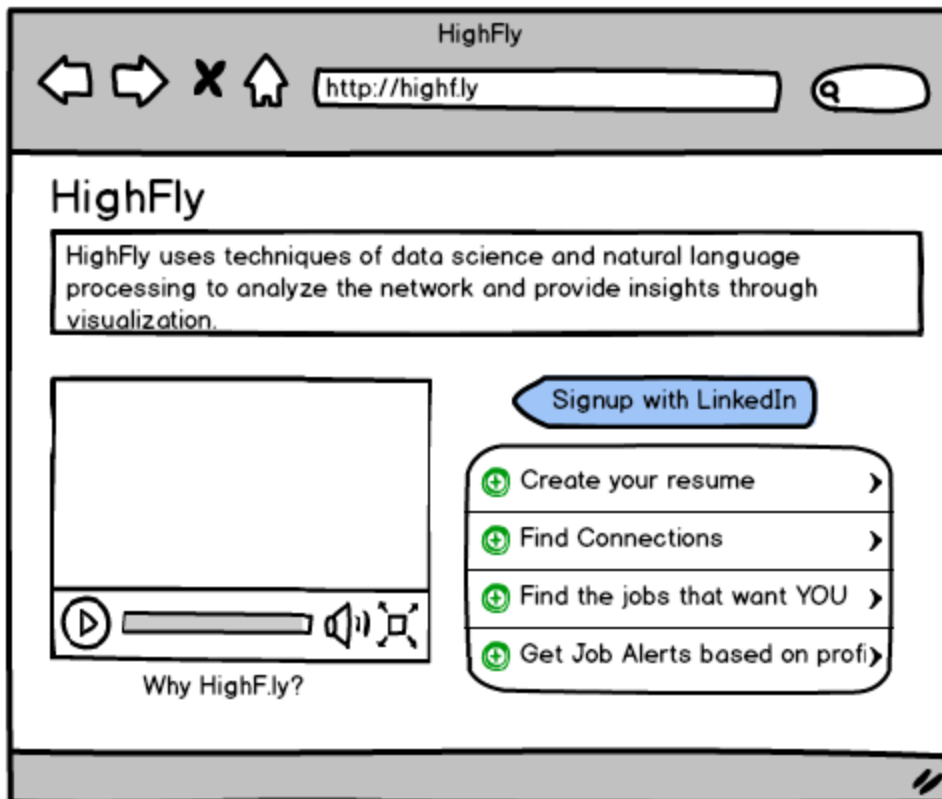
10. Job Alerts

There will be an option for daily job alerts based on user's expertise. Notification can be done via email. For a later version, we can think about a mobile client in which we can send push notification. This feature will expose the user to certain job posting from different sites which they might miss.

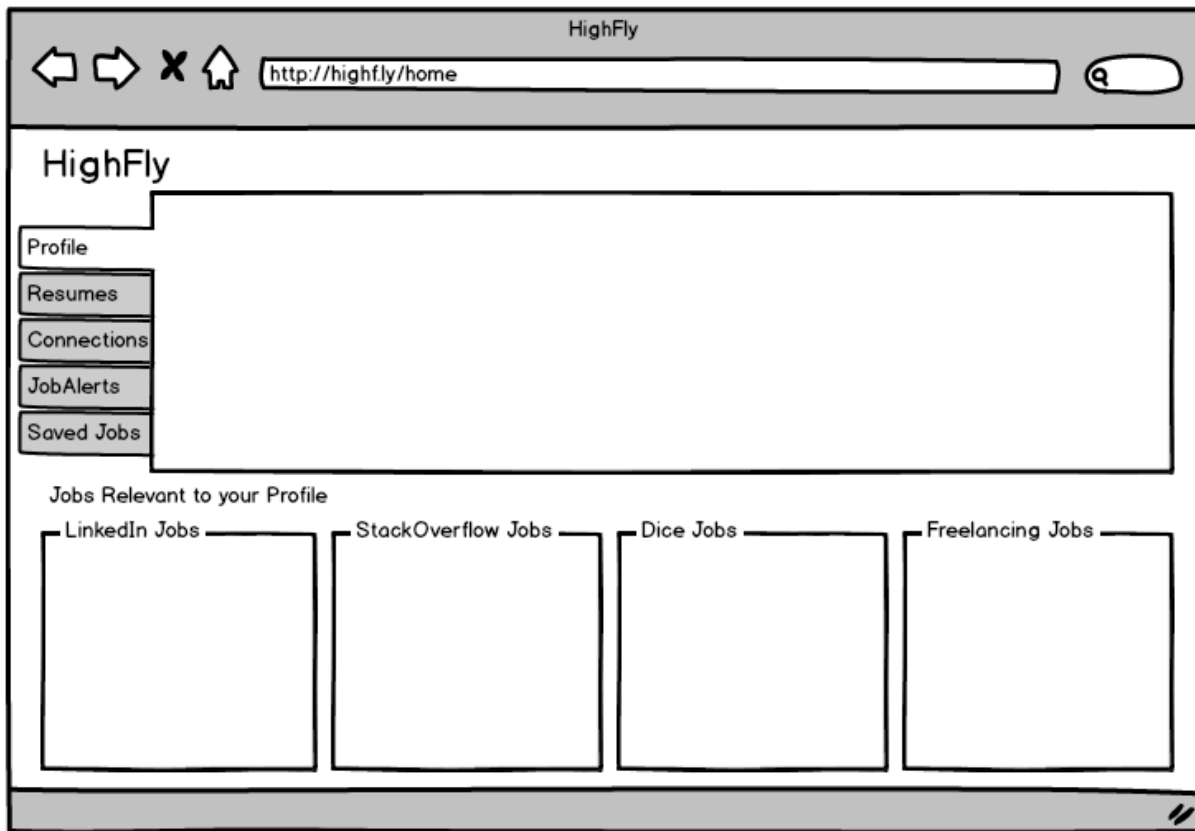
4.2 Visual Mockup

In this section we will try to shape up the idea through designing visual mockups according to the feature plans.

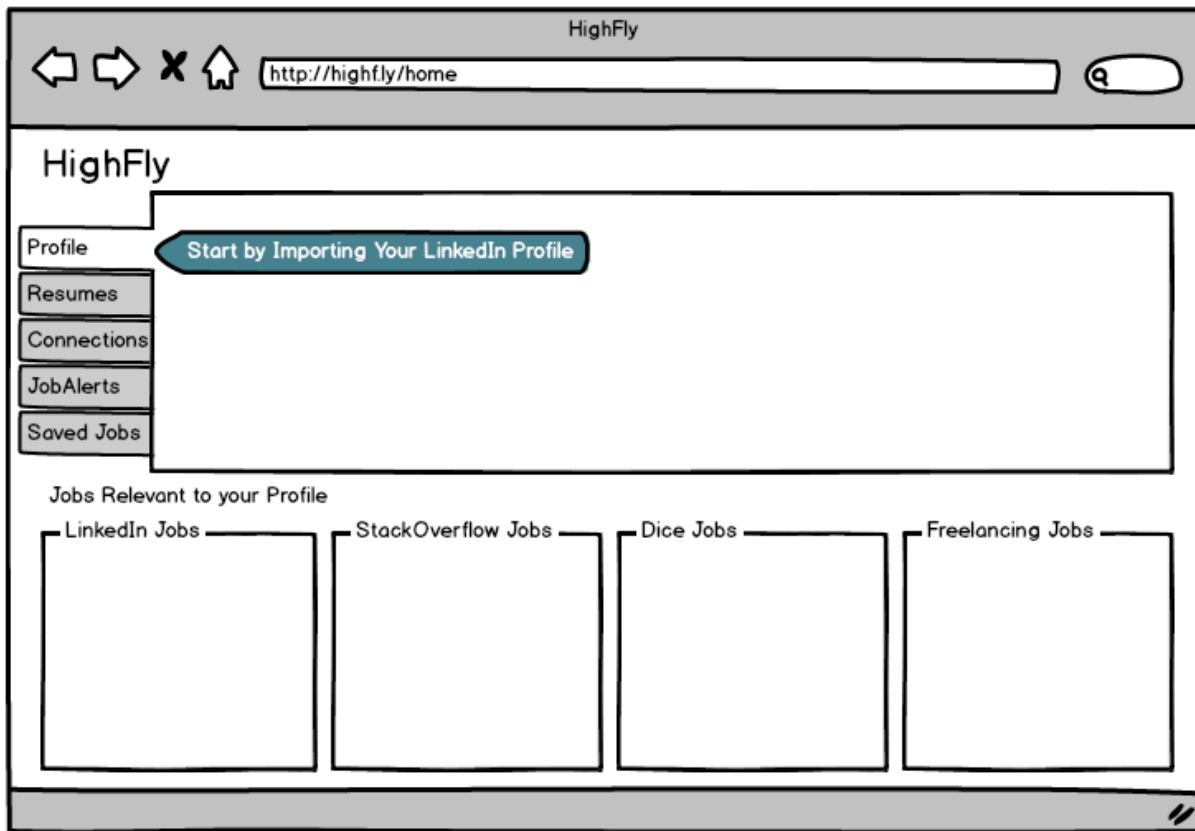
4.2.1 Application Landing Screen



4.2.1 Application Home Screen



4.2.1 Application Profile Screen [Before import]



4.2.1 Application Profile Screen [After import]

5 Development Plan

This section provides the current plan of implementation in terms of tools, technology and milestones.

5.1 Languages & Frameworks

This section describes different choices we have made so far in terms of technology (languages and frameworks). According to the feature sets defines so far, frameworks and libraries are chosen. These are subject to changes

5.1.1 Server Side

| | | |
|----------|--------|--|
| Language | Python | In order to utilize the large set of libraries available including natural language processing (like NLTK) and machine learning (like scikit-learn). |
|----------|--------|--|

| | | |
|-----------|-------------|---|
| Framework | Flask | <ol style="list-style-type: none"> 1. Minimal web framework built on Python. 2. Less boilerplate code. |
| Database | MySQL (TBD) | <p>This is not yet finalized.</p> <p>TO DO: Need to check integration of</p> |
| Hosting | Heroku | <ol style="list-style-type: none"> 1. Easy to configure and access. 2. Simple management. 3. Integration with version control. |

5.1.2 Client Side

| | | |
|-----------------------|---|--|
| Front End | HTML/CSS, Javascript (bootstrap and jQuery) | For responsive user interface to make the front end usable for devices of different screen sizes |
| Visualization Library | D3.js (Data Driven Documents) | <p>It is not finalized.</p> <p>TO DO: Need to check more libraries to create interactive graph visualization</p> |

5.1.3 Development

| | | |
|-----------------|------------------------|--|
| Version Control | Git | <ol style="list-style-type: none"> 1. Most developer friendly. 2. Integration with Heroku. |
| IDE | Eclipse, Atom, PyCharm | |

5.2 Development Plan

The milestones are defined according to the features described in weekly basis considering 8 hours per week coding session.

| Milestone | Date | Feature | Comments |
|-----------|------------|---|----------|
| 01 | 09/20/2014 | <ol style="list-style-type: none"> 1. Signup and Login using LinkedIn 2. Import User Profile Data 3. View Profile Data | |
| 02 | 09/27/2014 | <ol style="list-style-type: none"> 1. Import Connections 2. Cluster Connections | |
| 03 | 10/04/2014 | <ol style="list-style-type: none"> 1. Explore visualization libraries 2. Visualize annotated graphs and relationship with Connections | |

| | | | |
|----|------------|---|--|
| | | 3. Show different network metrics | |
| 04 | 10/18/2014 | <ol style="list-style-type: none"> 1. Search Companies 2. Search Jobs 3. Analyze Job requirements according to User Profile data | |
| 05 | 10/25/2014 | <ol style="list-style-type: none"> 1. Create Resume 2. Update resume from LinkedIn data | |
| 06 | 11/01/2014 | TBD | |

[1] <http://press.linkedin.com/about>

[2] <https://developer.linkedin.com/apis>

[3] <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/>