

**BILKENT UNIVERSITY**  
**ENGINEERING FACULTY**  
**DEPARTMENT OF COMPUTER ENGINEERING**

**CS 299**  
**SUMMER TRAINING REPORT**

**Mehmet Furkan Sahin**

**21201385**

**Performed at**  
**Memorial Sloan Kettering Cancer Center**

**Supervisor**

**Emek Demir**

**Start Date: 02.06.2015**

**End Date: 02.09.2015**

## **1. Introduction**

I completed my summer training program in Memorial Sloan Kettering Cancer Center. The company is mainly a cancer research center. MSKCC is putting effort to improve patient care, innovative research, and outstanding educational programs for 130 years. Today, MSKCC is one of the 45 National Cancer Institute–designated Comprehensive Cancer Centers, with state-of-the-art science flourishing side by side with clinical studies and treatment.<sup>1</sup> As an institution providing an environment that physicians, scientists, and engineers can work simultaneously on just one aim, they are very sure that in future, they will have an enormous contribution on the cure of cancer. Thus, they are able to do biomedical researches with the help of different domains such as computer science.

I, as a software engineering intern, worked in the Department of Computational Biology under the supervision of Emek Demir during my internship. I worked on 3 different projects, one of them was to mainly make me to understand the development environment, open source software development, and data visualization library Cytoscape JS. After the first project, I was able to work on a commercial project. I completed some tasks related to the user interface problem of SBGNViz<sup>2</sup>. After that, I was ready for a whole big and commercial project and took responsibility for the transformation of network tab of cBioPortal from a Flash based version to Cytoscape JS. The project is done and available in the official website<sup>3</sup>.

---

<sup>1</sup> <https://www.mskcc.org/about>

<sup>2</sup> <http://www.cs.bilkent.edu.tr/~ivis/SBGNViz.js/>

<sup>3</sup> <http://www.cbioportal.org>

## **2. Company and Supervisor Information**

John J. Astor, his wife Charlotte and a group of their friends founded Memorial Sloan Kettering Cancer Center in 1884 as New York Cancer Hospital on Manhattan's Upper West Side. In the 1940s, two former General Motors executives, Alfred P. Sloan and Charles F. Kettering, joined forces to establish the Sloan Kettering Institute (SKI), which has since become one of the nation's leading biomedical research institutions. Built adjacent to Memorial Hospital, SKI was formally dedicated in 1948.

In 1960, a new corporate entity — Memorial Sloan Kettering Cancer Center — was formed to coordinate and guide the overall policy for Memorial Hospital and the Sloan Kettering Institute, and in 1980 these entities were unified into a single institution, with a single president and CEO.

Over the years, they have continued to expand their outpatient facilities and services to meet the growing needs of their patients, physicians, and researchers.

My supervisor Emek Demir was a manager of the team Pathway Commons. He took his B.S. from Bilkent University, Molecular Biology and Genetics, M.S. from Bilkent University, Computer Engineering. He is currently assistant professor in Oregon Health & Science University.

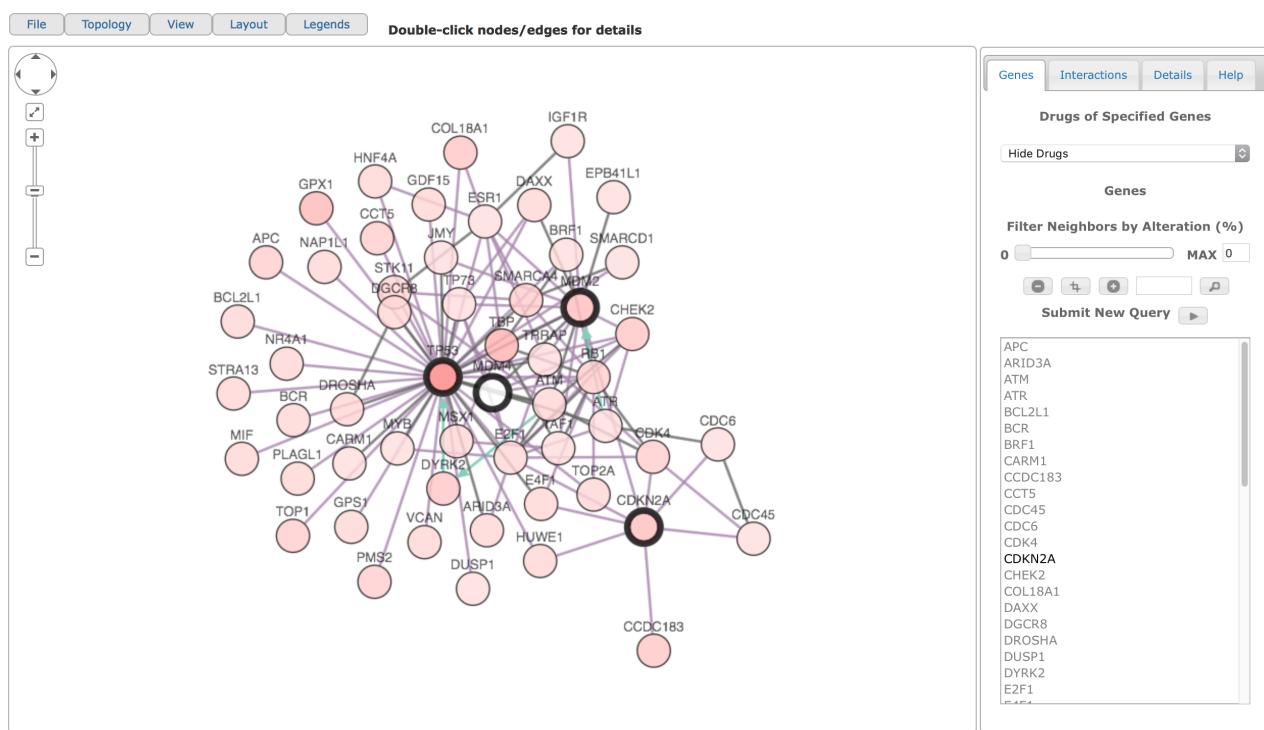
His mail address is demir@cbio.mskcc.org.

## **3. Work Done**

### **3.1. Main Project Information**

The main project that I worked on is the transformation of network tab of cBioPortal from Flash based version to Cytoscape JS. The tool mainly visualizes Genomics Data. The nodes represent mutated genes in the network and the edges represent their interactions. Color of the node shows the mutation ratio for that specific gene and the border of the node is drawn thicker if it is one of the

query nodes for that data set. The network data is fetched from the server and visualized. An example screenshot of the tool can be seen in Figure 1.

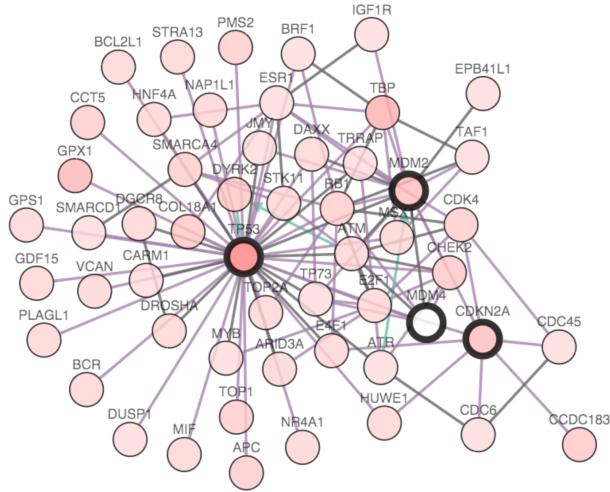


**Figure 1**

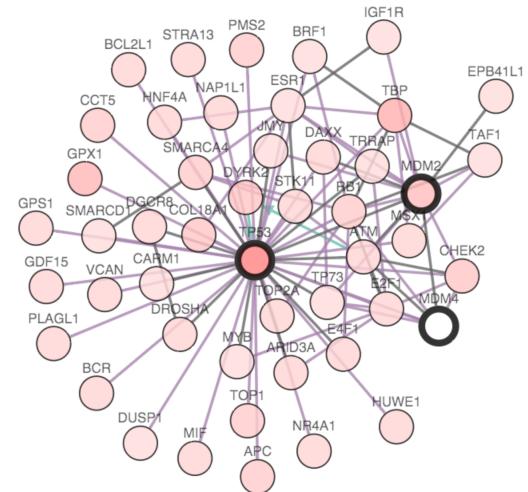
All of the menu, Genes, Interactions, Details and Help tabs were already implemented to be used with a Flash based network. I modified them so that they can be used with the JS version of the network. The tool is now deployed in [cbioportal.org](http://cbioportal.org) and used by the researchers.

### 3.1.1. Menu Functionalities

- The tool is able to save the current state of the network in png format. In this way, user is able to save his work and print when he makes an important observation.
- According to user needs, we provide functionality to hide some nodes and the edges linked to them. For example, you can see the difference between two networks in Figure 2.

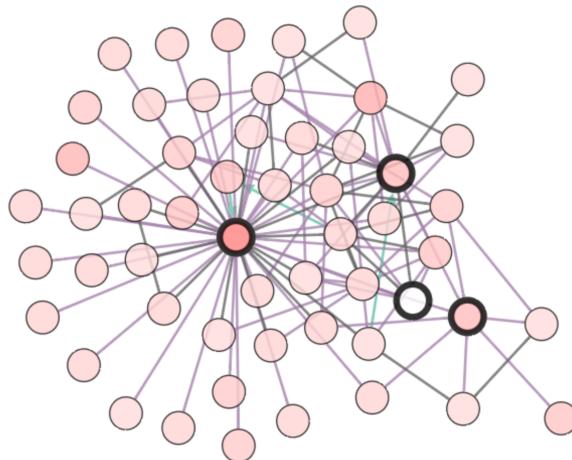


**Figure 2a**



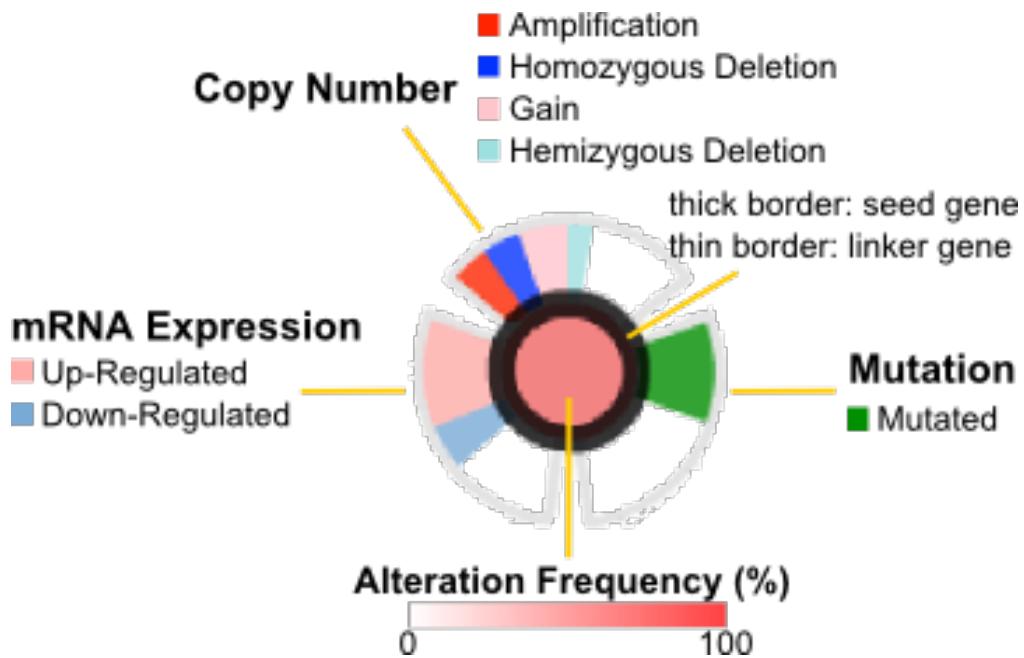
**Figure 2b**

- Id of each node is written on top of them and user might hide the labels if he wants as it can be seen from Figure 3.



**Figure 3**

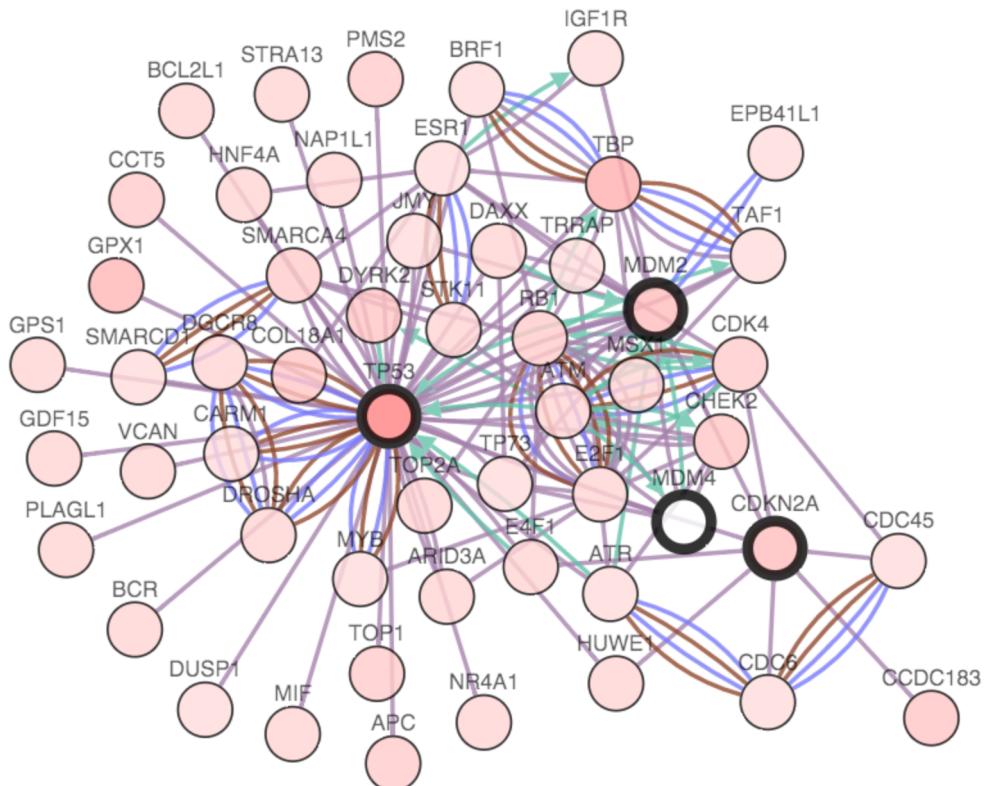
- In the tool, each node has a mutation profile. Mutation profile shows the details about the mutation history of that specific node. The details about the profile structure can be easily seen in Figure 4. Additionally, the same figure is provided in Legends to give an insight to the user about the tool. Also, the user has the ability to see all of the gene profiles in the same time by checking a menu item.



**Figure 4**

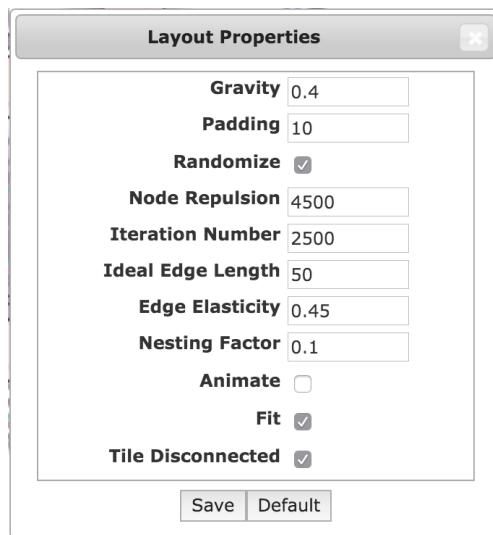
- In the tool, we automatically merge edges if they have the same source and target or crossed.

However, we provide user to see the merged edges, as you can observe from Figure 5.



**Figure 5**

- With the help of Pan-Zoom controller, user is able to travel around the network directly from user interface. Besides that tool, user is able to do the same gestures with the help of his mouse directly.
- The tool is able to apply the layout algorithm CoSE<sup>4</sup> on the networks and we provide a menu for the property changes of the algorithm as it is in Figure 6.

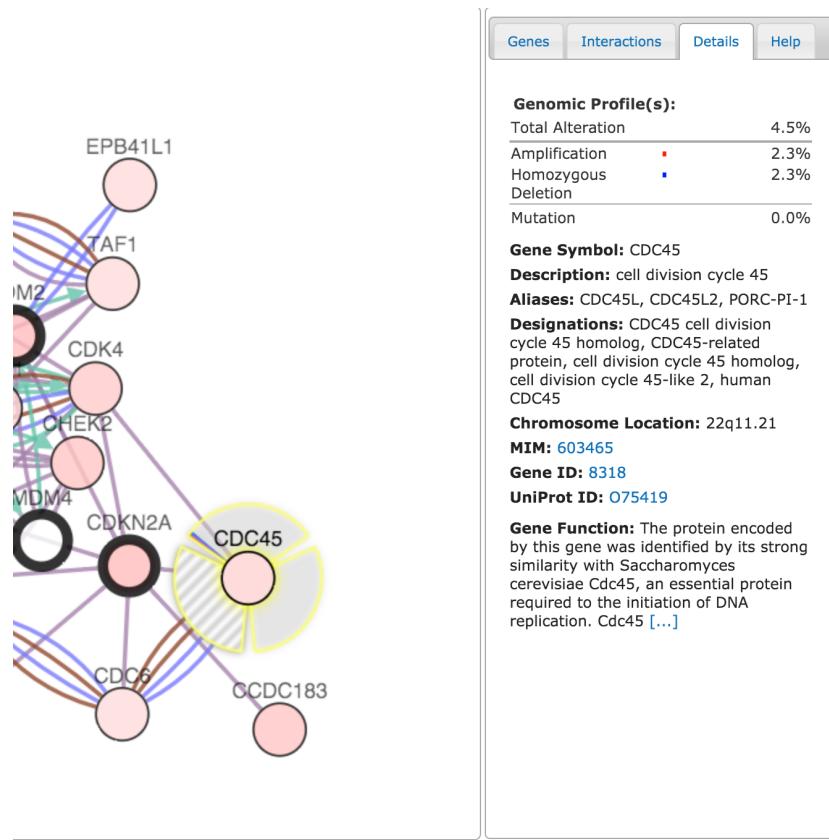


**Figure 6**

- On the right side of the tool, user is able to modify the current Query, choose some genes/ interactions, filter according to the alteration ratio or even see the details of that specific gene/ interaction. The detailed information is fetched automatically from the database when user double clicks on an edge or node as it can be seen from Figure 7.

---

<sup>4</sup> <http://www.cs.bilkent.edu.tr/~ivis/cose.html>



**Figure 7**

### 3.2. Significance and Motivation of Work

cBioPortal is created for researches in MSKCC. The tool is in usage for the aim to find a cure for cancer. The previous tool was created with Flash. Since the tool is a web browser based tool and web browsers are not supporting Flash anymore, thus, they needed the tool to be transformed to a new technology which is JavaScript. Since I was totally new to JavaScript, data visualization, and web development, my main motivation was to learn a whole new field. It was a big pleasure to learn a complete new domain of software engineering.

### 3.3. Detailed Description of Work

#### 3.3.1. Using New Tools and Technologies

**GitHub:** Since the organization uses git system. I had to learn and use it. I was totally new to the system but when I learnt how to use and benefit from it, I was appreciated. GitHub is used to

integrate the whole development process of the project. Github repository of the project is available in the following link; <https://github.com/cBioPortal/cbioportal>.

**WebStorm:** WebStorm is an IDE provided by JetBrains. After taking a static dataset sample, I created the whole visualization tool with its design specifications such as color, size, shape, border thickness, transparency, and edge attributes in WebStorm. The tool provides a very useful environment to work with html, JavaScript and Cytoscape JS library.

**Cytoscape.js:** It is a graph theory library for analysis and visualization. It is compatible with CommonJS/Node.js, AMD/Require.js, jQuery, npm, Bower, spm, jspm, Meteor/Atmosphere, and plain JS/JavaScript. Cytoscape.js supports many different graph theory use-cases. It supports directed graphs, undirected graphs, mixed graphs, loops, multi-graphs, compound graphs (a type of hyper-graph), and so on.<sup>5</sup>

**Google Chrome:** As a web browser, it provides a very useful debugging unit for web applications. Chrome is used in both development phase and then the testing phase with some other browser such as IE 10, Mozilla Firefox, and Opera.

### 3.3.2. Self-Training

I started working with a warm-up project which is called iVis-GVT<sup>6</sup> <sup>7</sup>. iVis-GVT was a simple project that I tried to create a Layout Algorithm visualizer. I used Cytoscape.js , Bootstrap, and html. A general overview of the application can be seen in Figure 8. The main aim on this project

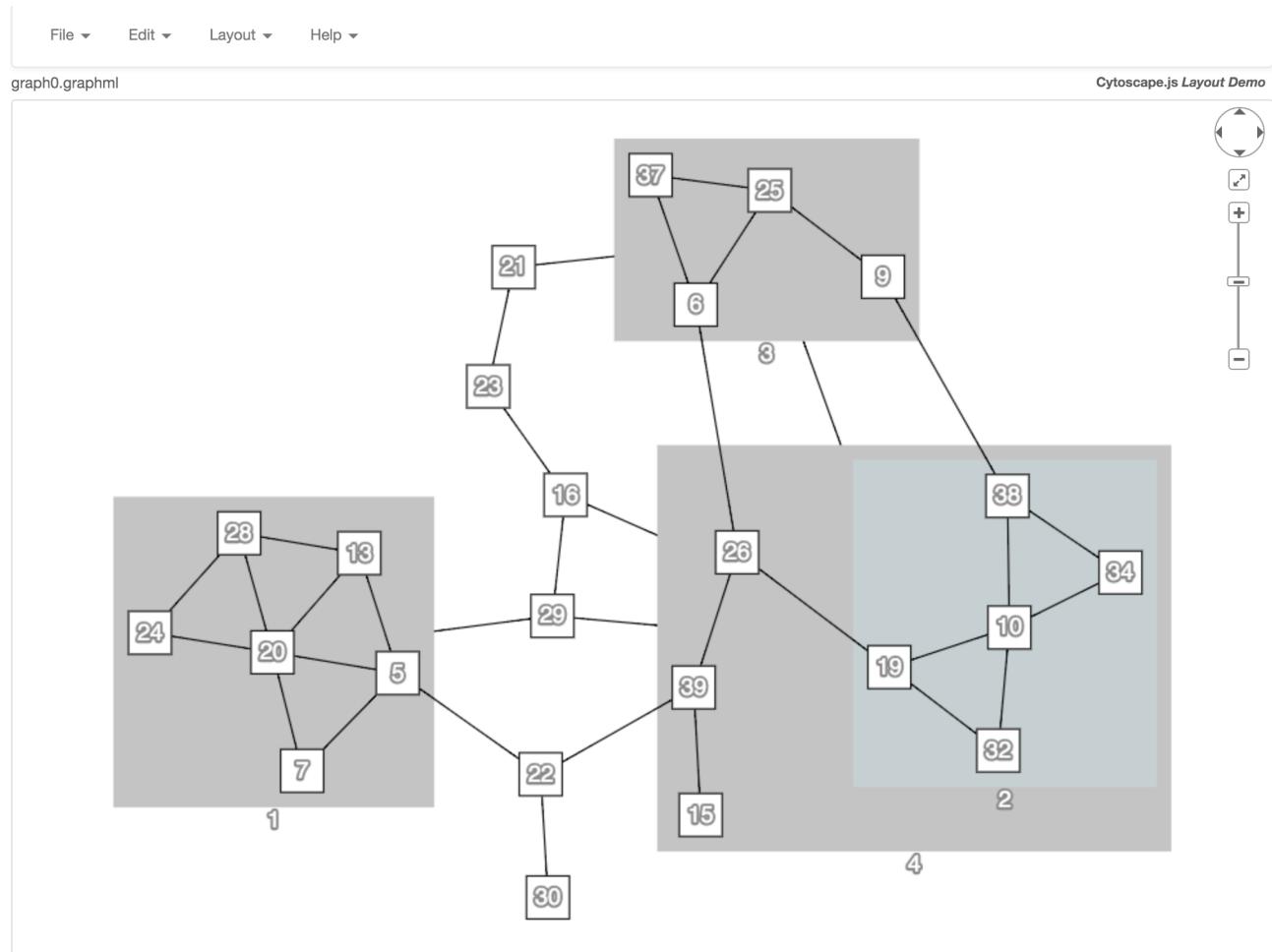
---

<sup>5</sup> <http://js.cytoscape.org>

<sup>6</sup> <https://github.com/furkansahin/iVis-GVT>

<sup>7</sup> [http://www.cs.bilkent.edu.tr/~ivis/Cytoscape\\_LayoutDemo/](http://www.cs.bilkent.edu.tr/~ivis/Cytoscape_LayoutDemo/)

was to make me more familiar with the environment and libraries that I were going to use during the main project. Thus, during this project, I learnt Cytoscape.js by the help of its documentation.<sup>8</sup>



**Figure 8**

The menu of the tool is created with Bootstrap. Thus, I also learnt how to integrate Bootstrap with html and create responsive and user friendly menus.

Since I had learnt html before my internship, I did not have any difficulty on creating a web page with html.

Thus, I worked on JavaScript libraries and web development on the warm-up project, I learnt how to integrate different web technologies and how to reach information about the libraries that I am working with.

<sup>8</sup> <http://js.cytoscape.org/>

After the implementation and testing stages of iVis-GVT, I took my internship project that was the transformation of the network tab of cBioPortal from a Flash based version to Cytoscape.js supported Javascript version. Since I was already learnt how to use Cytoscape.js and integrate it with html, I was able to create a design in my head but the main functionalities of the network was implemented in flash. I had to read some flash code and understand the functionality to implement them in the same way as they are in the Flash based version, since the bioinformatics researchers were very conservative about the current situation of the tool, I must not have been making any change in the functionality and user interface of the tool. The functions that I have to implement were generally menu related ones, because they were implemented to be working with a flash based network. However, thanks to my previous experiences on programming languages, I did not have any difficulty with reading flash.

### **3.4. Knowledge and Skills from School**

CS 202 Fundamental Structures of Computer Science II: I was able to understand the logic behind the graph theory algorithms.

CS 315 Programming Languages: With the help of that specific course, I was able to understand Flash easily and follow the syntax for the functions that I had to write in JavaScript.

### **3.5. Engineering Problems**

Before starting to my internship, I had coded in Python for a very short time. Thus, as a dynamic language, JavaScript was a new concept to me. For each of the bugs or errors in the code, I had to run the code and debug it in run time. Thus, it was hard for me to get used to code-debug cycle of JavaScript. However, I got used to the system by using developer tools of Chrome and WebStorm. Firstly, I tried to use WebStorm, but Chrome was better for me, since it provides both the web page in that time and a debugging environment.

After setting up the environment and starting to write the code, I had a problem on the sizes of a menu div. I was managing the website on css classes, but this time I was not able to detect which css class makes the size of the div less than I wanted. Also, this time Chrome debug environment was also not giving me the class hierarchy properly so that I can see the conflicts. In the class hierarchy that Chrome gives, everything was perfect and there shouldn't have been any problem at all. After almost 4 days of examination of the code, I came up with the solution. It was because of one of the Bootstrap classes. The class was changing the value that I decided and that class was also not visible in the Chrome class hierarchy. Actually, the main problem was that Bootstrap provides a user interface that user inputs the variables and dimensions he wants and the package is directly prepared for downloading and user starts to use it. In this phase, I put one of the values by mistake to the package and it made me to waste 4 days.

### **3.6. Work Environment and Team Work**

We, as a group of 3, were sharing the same office -me, and one more new developer like me and our supervisor-. Since we were very close in the office and working on very close projects, it was a big opportunity for me to get familiar with all of the new technologies. I was able to ask any question in any time. Also, my supervisor was the previous developer of cBioPortal's Flash based network tool, so, he was a huge information resource for me. Since, I had to read Flash code to understand the extra functionalities that I had to implement, I quickly got the information that I need in the times that I got stuck. He was also working on Cytoscape.js, thus he was able to answer my questions related to the library, too.

We all had our own projects, but still we were working on very similar tools. In such an environment, information sharing happens very quickly and questions get their answers with the same pace. Even if the guy that you asked the question does not know the exact answer, he knows where to get the right answer. This environment provides a big time and engineering power saving.

### **3.7. Professional Issues and Work-Related Ethical Issues**

We did not have a professional issue but one of the most important things that I had to get used to was to write header on the code properly. Since all of the projects of the company are open source and available in GitHub, they really care about the personal rights on the code. Each coder that makes a change on the code or writes a complete new class have to write his name on the proper places and license his code properly according to the general license policy of the company. Since they really care about the toil of each developer, they list his name on appropriate pages. For instance, you can find my name on the about page<sup>9</sup> of cBioPortal.

### **3.8. Multi disciplinary work**

Since we, as a group, were creating tools for a cancer research center, we had to communicate with the researchers and respond to their needs. During the development period, there are always 3-4 development feedback cycles. After each milestone in the application, it should be sent to the researchers and ask if it responds to their needs. Thus, this required me to write a couple of e-mails in an appropriate format. I learnt how to tell the tool that I created in a non-technical format.

### **3.9. Impact of Engineering Solutions**

I had chance to observe that engineering is used in very crucial parts of our lives. It is one of the most satisfactory things about my internship that I created a tool that is used by researchers to find a cure for cancer.

---

<sup>9</sup> [http://www.cbioportal.org/about\\_us.jsp](http://www.cbioportal.org/about_us.jsp)

#### **4. Conclusion**

During this internship period, I worked on mainly 2 projects that one of them warmed me up for the main project. Actually, during that first project, all of the library dependency, JavaScript based problems that I can have were solved. After this first project, the second and significant one was simply requiring me to apply the knowledge that I gained from the first one. During that period, the education that I got in Bilkent supported me remarkably, since I had to get used to be familiar with new Programming Languages quickly. Additionally, I saw the benefit of taking the education in english, since we documented everything in english and I worked on an english based development environment during the internship. After all of that experiences, I am proud of myself that one of the tools that I created is in usage of the benefit of humanity -especially it is used to find a cure for cancer-.