Eric Cherin
echerin@stevens.edu faster response send to eric.cherin@gmail.com
732 890 4421
Github ericcherin

8/29/14 I decided to join Dylan's group. A lot of these words that he uses to describe his project are not familiar. I need to brush up on vocab. I received a graph algorithm book, various graph algorithms, and lots of links. Dr.Ganesan has cluster of nodes. We may use Accumulo. Look at d4m as well. I am going to get access to the team's Google drive.

8/30/14 Met with Dylan to discuss project. He introduced me to project and workflow. We will be using a lot of different software for our project such as github and todoist.

9/2/14 I agree that we need to make this project further to the right. We cannot just research stuff. We need to create a sellable product in order to fulfill tg and senior design requirements and approach. Dylan's argument is that we can market our results as a package that big data companies would use.

9/4/14 Met with di ren, but Andres and Edward didn't show up the meeting. We need to find senior design group members. Our team has a google drive, todoist webpage, and a github group. I am thinking that Xin Li may still need a group.

9/5/14 Xin li and Hefei Yang said they can be a part of the group! I emailed Dylan that they wanted to join and we plan on having a meeting tomorrow.

9/6/14 We have to do weekly effectiveness reports and tg stuff. We have 10 people and an advisor working on this project.

9/8/14 Dylan sent this to me:

Hi Eric,

Yes, it's a good book but easy to get overwhelmed when flipping through. **Read smartly**. Better to understand one thing thoroughly than to have vague notions about many things. I advise reading with pencil and paper at your side, so that you can try things on sample graphs.

Chapter 1 is intro material and chapter 2 is on general principles of linear algebra. They are both nice to know but not strictly necessary. Skip them if you want to go directly to your first graph algorithm.

Chapter 3 and 4.3 are tutorials on first graph algorithms. You wlll learn about (strongly) connected components and shortest paths in two different ways, and some other related constructs. These are good places to start.

Chapter 5 starts getting serious with Bellman-Ford shortest paths, all-pairs shortest paths and minimum spanning tree.

Future chapter I haven't looked at much yet. We will not use all the material in them, but probably some.

When you read the stuff about *semirings*, the book is referring to replacing the '+' and 'x' operations of matrix multiplication with something else, such as 'min' and '+'.

~Dylan

9/9/14 I looked at computing with gpus in java. We have to decide which programing language to use. We are thinking about using OpenCL over Cuda. We can write in higher language programs if we like, but then someone will have to convert it to a lower level language later on. I researching D4m as well. I made an environment for opencl and wrote hello world.

9/14/14 Our project will be open source. I looked at Dylan's code for strongly connected components. I know the algorithm, but there one line that I do not understand in respect to syntax. I fully understand the strongly connected component algorithm, however.

9/15/14 I do not think we should meet up to do the TG lab. We can simply do it over google drive. Everyone agreed that it was not efficient to meet just to discuss a tg lab. I saved the team a lot of time.

9/18/14 Our roles in the project are clear. I will assist in TG stuff, Website design, and graph algorithm prototyping. I need to read up on the related work section in google drive. I also downloaded netbeans ide.

9/23/14 We now have a license to the parallel computing toolboxin Matlab. I need to understand what we are doing with CUDA. I do not have the appropriate gpu to run Matlab parallel processing on my machine. I need to learn how to use linux and connect to servers in order to run matlab.

9/25/14 I got a review of Accumulo. Now I have to learn how to navigate and connect to the server. I am new to linux operating systems.

9/30/14 Dr.Ganesan recommends we should try to have a more concrete focus such as biological application. I agree that we should focus on something more concrete. I looked through the project proposal requirements.

9/30/14 I feel like our group is too split. We have about half people working on biological application stuff while the rest of us are working on graph algorithms. Dylan is fine with splitting everything up, but I feel like we need to work on the same goal as a team.

10/2/14 I looked at the hmmer algorithm. I researched the different types of databases. I am still not clear with the iterators, I need to read into that more.

10/3/14 I looked at more algorithms for graph stuff. I looked at graph partitioning and non negative matrix factorization.

10/5/14 I met up with Jaroor and we discussed algorithms in the library. We sketched out exactly what we needed to do with the prim graph algorithm. I learned how to do basic commands and open a text processor in linux.

10/9/14 I am having slight troubles with my github. I had to install Microsoft's shiny version of git, which kind of worked well but was very slow. We finally stopped working on the graph algorithm. I am tasked with learning how to use JNI. It looks kind of scary so I will look at this later. We do not immediately need to use this anyways.

10/13/14 I worked on the proposal a lot today. I am trying to install accumulo on windows. I looked at tutorials for accumulo as well. I have midterms and lots of work coming up in the next two weeks, so I probably will not get a lot done for this project.

10/17/14 Dylan got accumulo running on the server. However, there is a whole process and lots of new linux commands that I need to learn. I will have to learn this after midterms.

10/20/14 I volunteered to be the proposal leader. My role is to add or get people to add new things to the proposal. I will also lead the group to make the PowerPoint.

10/26/14 I looked at this and read relevant papers.

1. MSV filter
2. Viterbi filter
3. Forward score

10/28/14 I read up on the JNI and started looking through tutorial code. My past experience in C++ has been very basic coding syntax. There are a lot of words and syntax that I did not know. This is a very time heavy learning process. I will be better suited to work on Accumulo with the skills that I have.

10/29/14 I installed intellij, WinSCP, and tried to install tortoisegit again. I also cloned the accumuloiter directory.

11/3/14 There is a really annoying "feature" on intellij installed on my machine. For some reason the blinking line where you input text jump to after the parenthesis: ()| it is confusing because if you click like this: (|) the input will be inserted after the parenthesis. I may have to install a different IDE because this has happened in the past and it wasted so much time.

11/6/14 We worked on the presentation PowerPoint. I have been programming using accumulo as well.

To be continued