# APPM 2360 Lab #3: Class President

Alex Mault, Evan Mori, John Dunn

December 6, 2013

# 0 Introduction

Last year, the elections for class president were rigged at the high school. The elections were based off of popularity and were held using a traditional election method. Obviously high school students cannot be trusted; therefore, the principal of the high school has decided to utilize Facebook and the Page Rank algorithm used by Google in order to determine class presidents from now on. This new method will allow the class presidents to still be chosen based on popularity, but without the use of individual voting. Instead, the number of Facebook friends a given candidate has will serve as a gauge for their popularity. Using the idea of eigenvector centrality the importance of each friend will also be taken into account. This will prevent the system from being easily rigged by the creation and addition of fake Facebook accounts.

The principal does not know how to execute this new method by himself, instead he has hired an outside group of programmers (us) to set up the programs necessary for the elections. In order to make sure that this group will be able to effectively execute the program, the principal has required that a warm-up exercise be completed in order to show the proficiency necessary to execute the desired program.

# 1 What kind of matrix has the following property?

$$M_[i,j] = M[j,i]$$

The property in question defines a symmetric matrix.

# 2 Is the friendship adjacency matrix symmetric?

Yes, the friendship adjacency matrix is symmetric. It is symmetric because the friendship is undirected, meaning that if Person I is friends with Person J (M[I,J]) then Person J is also friends with person I (M[J,I]).

From these warm-up exercises, the programmers showed that the friendship adjacency matrix (the matrix that will be utilized to show popularity of the friends of each candidate) is a symmetric matrix. It is no surprise that the friendship adjacency matrix is symmetric because Facebook friendships are mutual and undirected connections. This means that if a candidate is friends with someone, that person is also friends with the candidate.

# 3 When Google uses the Page Rank algorithm, it creates a graph of the network of websites, with web links serving as the connections. Is this graph directed or undirected? Why?

The Page Rank algorithm is defined, in general, for directed graphs. This is because clicking on a web link from a Google search typically directs the user to the desired web page and a link on the web page does not usually direct the user back to Google.

# 4 Construct an adjacency matrix from the following network. Does it have the property described in problem #1?

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 2 | 0 | 0 | 1 | 0 | 1 | 1 |
| 3 | 0 | 0 | 0 | 1 | 0 | 1 |
| 4 | 0 | 0 | 0 | 0 | 1 | 0 |
| 5 | 1 | 1 | 0 | 1 | 0 | 0 |
| 7 | 0 | 1 | 1 | 0 | 0 | 0 |

Table 1: The adjacency matrix is not symmetric

# 5 Give the centrality scores of these books. What does this tell you about the books?

| Books | Centrality Score |
|---|---|
| 1 | 6 |
| 2 | 7 |
| 3 | 3 |
| 4 | 3 |
| 5 | 7 |
| 6 | 5 |

If a person were to count the 1s in each individual row, the number of books each corresponding book cites would be achieved. These scores are as follows: Book 1 cites 3 books, Book 2 cites 3 books, Book 3 cites 2 books, Book 4 cites 1 book, Book 5 cites 3 books, and Book 6 cites 2 books. In this case, 3 books is the largest number of citations a single book contains. The centrality score is the summation of the number of citations each sited book contains. A higher centrality score indicates that there may be more information in the book than other books with lower scores.

# 6 What is the rate of convergence for this problem?

To find the rate of convergence, we have to calculate the largest eigenvector using the equation below where $b_0$ is a random vector.

$$b_{k+1} = \frac{Ab_k}{||b_k||} \tag{1}$$

If $b(i)$ is the $i^{th}$ element of the vector $b$ of length $n$,

$$||b|| = \sqrt{\sum_{i=1}^{n} b(i)^2} \tag{2}$$

We can calculate this using the matlab *norm* command.

This process has many iterations. For each iteration $k$, if $b^*$ is the true eigenvector corresponding to the leading eigenvalue, the error at iteration k is given by:

$$e_k = ||b_k - b^*|| \tag{3}$$

The rate of convergence is the ratio:

$$r = \frac{e_{k+1}}{e_k} \tag{4}$$

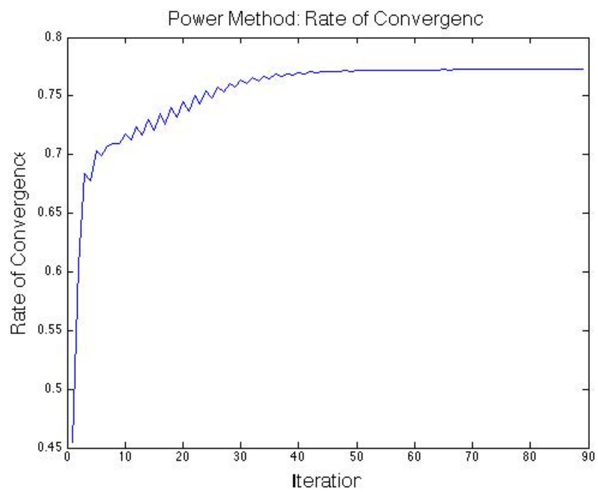


Figure 1: In the election example, the rate of convergence approaches 0.7726.

# 7 Plot the error, $e_k$, against k.

Obviously, with the power method, the error will get smaller and smaller each time, approaching zero in the last iterations.

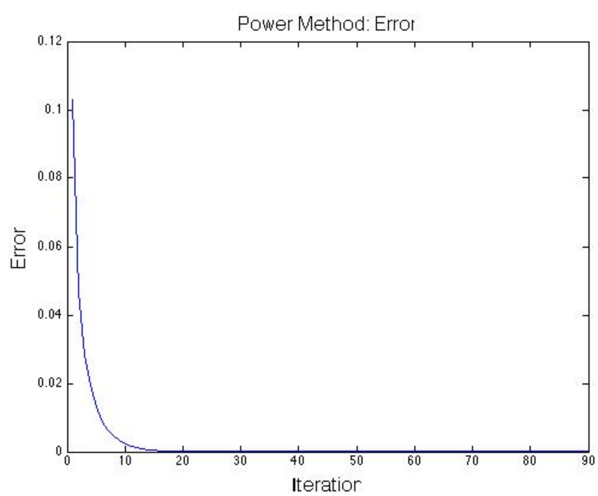


Figure 2: Eq(3) establishes the error against k.

# 8 Rank the students P1-P15, as well as the candidates, with scores based on the Facebook adjacency.

To rank the students based on there centrality, we first had to construct a friendship adjacency matrix. Students P1-P30 are represented as the first 30 rows and columns. The last three rows are Mary Fred and Veronica respectively.

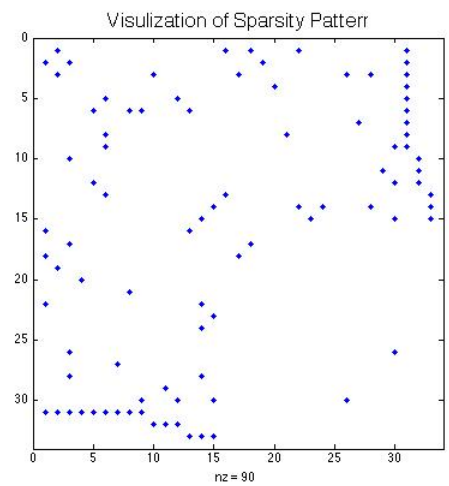


Figure 3: visualization of connections between friends. Each friendship connection is indicated by a dot.

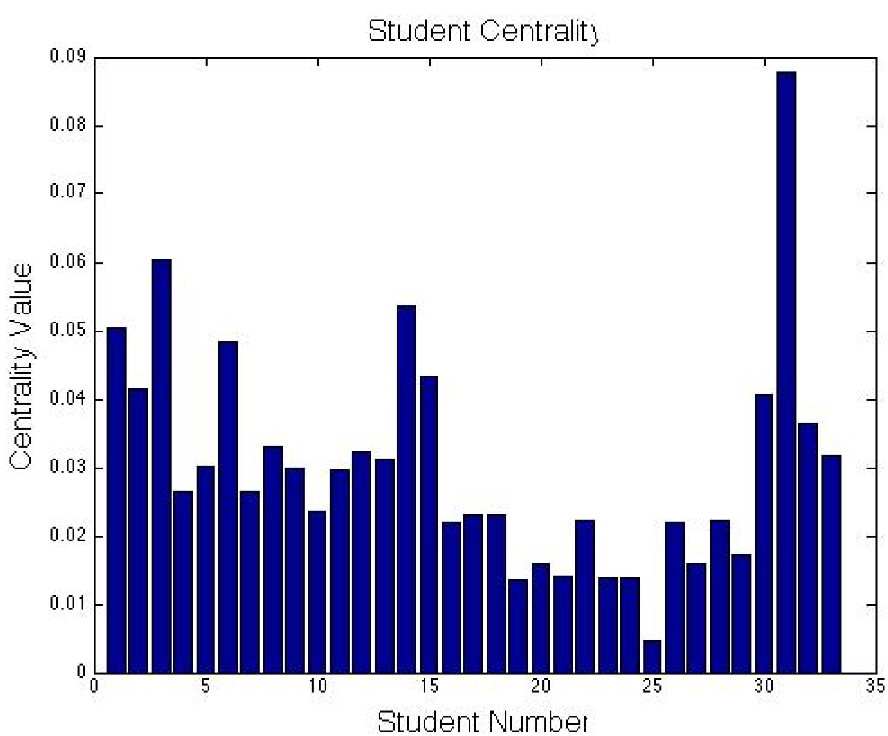


Figure 4

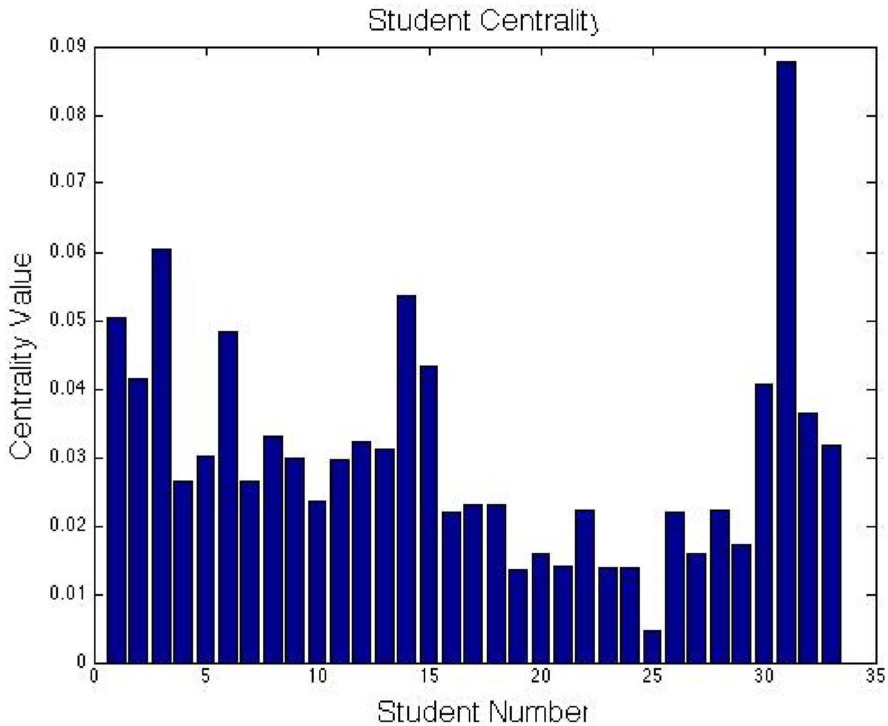


Figure 5

## 9 How could someone cheat to make their score higher with this system?

## 10 What are some different flaws in this system of ranking the students' popularity? How could it be changed to correct that problem?

## 11 Would you have picked the same student to be class president? Why or why not? Are there other factors that you would take into consideration to make the system more balanced?

## 12 What are some examples of other networks that could be modeled using an adjacency matrix?

5. Give the centrality scores of these books. What does this tell you about the books? a) Counting the 1's in rows gives how many each book cites Book 1: 3 *** Highest *** 2: 3 *** Highest *** 3: 2 4: 1 5: 3 *** Highest *** 6: 2 Counting the 1's in columns gives how many each book is cited Book 1: 1 2: 2 3: 2 4: 3 5: 3 *** Highest *** 6: 3 *** Highest *** If have to combine them for some reason. Book 1: 4 2: 5 3: 4 4: 4 5: 6 *** Highest *** 6: 5

Mary Friends with P1, P2, P3, P4, P5, P6, P7, P8, P9. Fred Friends with P10, P11, P12. Veronica Friends with P13, P14, P15.

| Centrality | Student |
|---|---|
| 0.0504 | 1 |
| 0.0415 | 2 |
| 0.0603 | 3 |
| 0.0264 | 4 |
| 0.0302 | 5 |
| 0.0481 | 6 |
| 0.0264 | 7 |
| 0.0330 | 8 |
| 0.0297 | 9 |
| 0.0235 | 10 |
| 0.0296 | 11 |
| 0.0321 | 12 |
| 0.0312 | 13 |
| 0.0535 | 14 |
| 0.0431 | 15 |
| 0.0876 | Mary |
| 0.0363 | Fred |
| 0.0318 | Veronica |

Figure 6: cats