INSTITUT FÜR INFORMATIK
der Ludwig-Maximilians-Universität München

Bachelorarbeit

Eigenvalue-based Bipartite Ranking

Stephan Link

Aufgabensteller
und Betreuer: Prof. Dr. François Bry, Christoph Wieser
Abgabedatum: 18. November 2011

Danksagung

Ein besonderer Dank gilt Prof. Dr. François Bry, der die Idee hatte auf der diese Arbeit basiert.
Besonders danken möchte ich auch meinem Betreuer Christoph Wieser, der mich bei meiner Arbeit sowohl fachlich als auch menschlich unterstützt hat.
Bedanken möchte ich mich auch bei meinen Eltern Elisabeth und Helmut Link, die in der anstrengenden Zeit des Schreibens immer verständnisvoll waren und mich moralisch unterstützt haben.
Besonders danken möchte ich auch meinen Kommilitonen Barry Norman und Jonathan Staude, die mich in vielen fachlichen Fragen auf alternative Sichtweisen und gute Ideen brachten.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst habe und keine anderen als die angegebenen Hilfsmittel verwendet habe.

München, den 18. November 2011

Stephan Link

Abstract

Rankings have a high relevance in many different domains, such as science, sports or education. According to the needs of applications and their data several ranking techniques were developed. In this thesis, we introduce a ranking technique called *bipartite ranking*, that is based on PageRank [8]. *Bipartite ranking* is suitable to rank scenarios, that can be represented as a bipartite graph what is among others the case for Rankings with feedback ranking in which those who were rated rate the person who rated them. We show, that the ranking results of *bipartite ranking* have a higher objectiveness towards defined criteria than existing ranking techniques. And we give examples for real situations, where *bipartite ranking* could be used in.

Zusammenfassung

Rankings stellen in fast allen Bereichen des Alltags, der Wissenschaft und Wirtschaft eine wichtige Entscheidungsgrundlage dar. In dieser Arbeit stellen wir das auf PageRank [8] basierende Rankingverfahren *bipartite Ranking* vor, dass speziell für Szenarien, die als bipartiter Graph dargestellt werden können entwickelt wurde. Szenarien, die als bipartiter Graph dargestellt werden können sind unter anderem Rankings mit Feedback Rankings, also Rankings die Szenarien in denen diejenigen, die bewertet wurden diejenigen bewerten, die sie bewertet haben. Wir zeigen, dass die Rankingergebnisse mit diesem Verfahren im Vergleich zu bestehenden Verfahren hinsichtlich vorher definierter Kriterien aussagekräftiger sind und geben Beispiele für reale Situationen in denen dieses Rankingverfahren benutzt werden kann.

Contents

Chapter 1

Introduction

Rankings have a high relevance in many different domains, such as science, sports or education. Rankings, that order a set of specific elements by predefined criteria belong to everyone's daily life. Rankings are not only used for illustration purposes in newspapers but are very often the basis of important decisions. Since rankings have a high importance, we believe that it is very promising to look at existing ranking techniques and find a way to increase their objectiveness.

In the beginning we were concentrating on a specific ranking: Grades for exams at universities, that can be interpreted as a ranking as well when ordering the students of a specific year by their average grades. In discussions with students we found out that many of them are not satisfied with the expressiveness of such rankings. They criticize that grades heavily depend on the lecturer and since students can organize on their own when they do an exam, that issue influences the ranking heavily.

In order to have both sides involved we were talking to a couple of lecturers as well. In the discussions we talked about lecturers rankings. These are rankings where students get the possibility to give feedback to their lecturers and grade them. Lecturers then get ordered by their average grade. The lecturers where criticizing an issue that is similar to what the students told us: The lecturers rankings are heavily influenced from bad ratings that come from students who were not able to pass the final exam.

We then concluded that the lecturers as well as the students are actually asking for a new ranking, where the influence of single ratings depend on who the ranking did. More precisely, the influence of a rating should depend on the raters reputation. If we now look at the students ranking and the lecturers ranking of a specific course of studies at a specific university (e.g. Computer Science at the LMU) we have the case that the lecturers in the lecturers ranking are the raters in the students ranking and vice versa for the students. Our objective in this thesis was then to create a ranking technique where we can give the two separated students and lectures ranking into as input and finally receive new students and lecturers rankings in which the weight (or influence) of single raters depends on the rater's reputation. The reputation then comes from the input rankings. More precisely: The highest rated lecturer's rating has the highest weight and vice versa for the students.

In the following sections, we introduce our ranking technique, that we call bipartite ranking. We will stay at the students and lecturers example, that we will introduce more detailed in chapter 3 before we finally describe the bipartite ranking technique. In chapter 4 we will evaluate bipartite ranking by comparing it to existing ranking techniques in specific scenarios. In chapter 6 we finally introduce two further applications for our ranking.

Chapter 2

Related Work

In this section we want to introduce work that is related to our topic. We will first have a look at the website www.meinprof.de which is a german portal that allows students to rate their lecturers. Furthermore, we will introduce the fundamentals of PageRank which is the ranking technique Bipartite ranking is mainly based on. In the last part of this section we will have a look at existing ranking techniques that a follow a similar approach but with different applications.

2.1 Lecturer Rankings on www.meinprof.de

The website www.meinprof.de offers students the possibility to rate their lecturers by giving them grades regarding eight specific criteria. Criteria are for instance:

1. Fairness
2. Support for the students
3. Comprehensibility
4. Humor

Students give one rating between one and six (the scale ranges from very good to very bad) for every criteria. The final score for a lecturer is the average out of all single grades. All criteria have the same weight.

In comparison to the evaluation that exists at the most universities the ratings are public and can be accessed by everyone. While the evaluation at university usually takes place during or after the last lesson but before the date of the exam, at www.myprof.de Students have the opportunity to grade their lecturers at any point of time.

Figure 2.1: Screenshot of www.meinprof.de

2.2 PageRank Fundamentals

This section introduces the fundamentals of PageRank as basis for our ranking technique. In [1], there are a few axioms presented, that are satisfied by PageRank.
Further important information dealing with the PageRank and properties of the damping factor, we found in [6], [2] and [7].
Because the bipartite ranking is a ranking technique for bipartite graphs, it is also promising to have a look at [9] where fundamentals of graph theory are introduced and explained.

PageRank is a website-ranking algorithm, that was developed by Larry Page and Sergey Brin [8]. PageRank interprets the World Wide Web as set of linked documents. The idea behind PageRank is to rank documents by their importance. For that, it assumes that a document, that is linked by an important document is more important than a document that is linked by a less important document.[3]
In the following short example, there is a set of four documents, that are partly linked to each other. Figure 2.2 shows a graph that represents a set

of linked documents. Each node represents one document. An edge between two nodes means, that the document that the edge starts from, contains a hyperlink to the document where the edge ends.

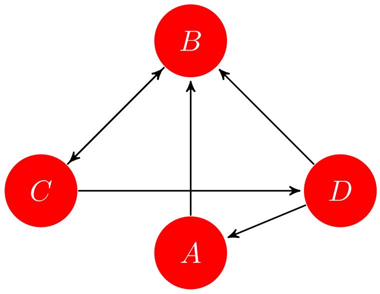

Figure 2.2: The link structure of 4 documents represented as a graph

As you can see in figure 2.2, all documents link document B. According to PageRank's assumption, document B is important. On the other hand, B itself links only to one document: document C. So document C should also be important. When computing the PageRank for that example, the solution vector looks like this:

$$\begin{pmatrix} 0.091068 \\ 0.363614 \\ 0.363500 \\ 0.181818 \end{pmatrix} \tag{2.1}$$

Figure 2.3: The solution vector of the PageRank computation of the example in Figure 2.2.

The first component of the vector refers to node A, the second to node B, the third to node C and the last one to node D. As we can see, the PageRank computation rates these documents very high, that we defined to be important. (B and C)

2.3 Existing Ranking Techniques Based on PageRank

We were not able to find a publication, where the computation of an eigenvalue is used to rank a bipartite graph. We therefore showed interest in ranking techniques, that are also based on PageRank, but not particularly for bipartite graphs.

FolkRank
One of these rankings, that is based on an adapted PageRank, is Folkrank (See [4]). Among others, Folkrank can be used in order to search in folksonomies. A folksonomy describes users, resources, tags and the user-based assignment of tags to resources [4]. And it can also be defined as undirected tripartite hypergraph.
The Folkrank algorithm uses a similar attempt like our bipartite PageRank, but the structure of the data, that FolkRank ranks differs from ours.

Chapter 3

Bipartite Ranking

3.1 Data Model

This section introduces the data model for our bipartite ranking and in particular the requirements of ratings:

1. Students do never rate other students, the same is valid for the lecturers. Because of that rule, the graph representation of our rating tables will always show a completely bipartite graph.

2. All students rate all lecturers and all lecturers rate all students. In practice it might happen, that single ratings are missing. In this case, one has to define a strategy for the completion of incomplete ranking tables. The task of the completion strategy is to guess the expected value. The expected value depends on the rating behaviour of the element that should have made the rating, as well as on the expected rating value for the element that should have received the rating. We assume, that missing ratings will inevitably influence the results of the whole ranking. However, in this thesis we defined, that there are no missing ratings.

3. The rating scale starts from 1 which means very bad and ends with 15 which means very good. This scale is also used at german high schools.

4. Both groups (e.g. lecturers and students) have to use the same rating scale.

In figure **??** you can see the rating table for an example with 3 students and 2 lecturers. The entries in the rating table show how a specific student ($S_i, i \in \{1, 2, 3\}$) rates a specific lecturer ($D_j, j \in \{1, 2\}$) and vice versa.

	S_1	S_2	S_3	D_1	D_2
S_1	0	0	0	15	8
S_2	0	0	0	5	12
S_3	0	0	0	5	12
D_1	15	4	4	0	0
D_2	12	12	12	0	0

Figure 3.1: Rating table with 3 students and 2 lecturers

This matrix can be represented as a complete bipartite graph. You can see the graph representation in figure 3.2.

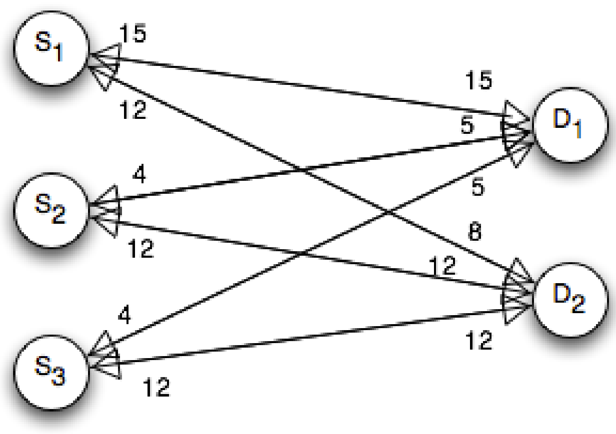

Figure 3.2: Graph representation of the rating table in figure 3.1

3.2 Use Cases

In this section, we want to introduce a ranking technique, that might be better than the ranking techniques, that are usually used. We suppose, that the widest spread ranking technique is simply to determine the average of all received ratings, for every student or lecturer in the rating table. Therefore, in section 3.2.1, we want to show an example for a usual average ranking for the students and lecturers application. While taking a closer look at the results of the average ranking, we want to explain, why we believe that there might a better ranking technique than the average ranking. In section 3.2.2, we will introduce PageRank as ranking technique for the students and lecturers application. In the last section of this chapter, we will finally introduce our bipartite attempt. This attempt uses PageRank, as the attempt in the previous section, but is exclusively suitable for bipartite graphs.

3.2.1 Average Ranking

In this section, we want to make a simple average ranking. You can take a look at the data that we are going to use in figure 3.3.

	S_1	S_2	S_3	D_1	D_2
S_1	0	0	0	2	12
S_2	0	0	0	5	12
S_3	0	0	0	5	12
D_1	2	12	12	0	0
D_2	15	10	10	0	0

Figure 3.3: Rating table for 3 students and 2 lecturers

We are going to make make two different rankings. One for the students, and the other one for the lecturers. The higher the average rating, the better the corresponding element gets ranked.

	average rating
S_1	8.5
S_2	11
S_3	11

Figure 3.4: Average ranking for students

	average rating
D_1	4
D_2	12

Figure 3.5: Average ranking for lecturers

In order to generate data for the example above, we defined types of students, and types of lecturers. We have a good lecturer (D_2), and a bad lecturer (D_1). The students S_2 and S_3 are homogenous, and gave and got the same ratings. Furthermore, there is a student (S_1), that has the lowest average rating of all students. He was rated very bad by the bad lecturer (2), but rated very good by the good lecturer (15). On the other hand, S_2 as well as S_3 got average ratings by both lecturers. If we make the assumption, that the good lecturer's ability to rate students is more convincing than the one of the bad lecturer, we come to the conclusion, that the results of the average ranking are not fair. It would be better, if the rating that came from the good lecturer would have a higher weight than the one of the bad lecturer.

This leads to PageRank, a ranking technique for linked documents. Among other things, PageRank can provide an individual weight for every rating.

3.2.2 Non-Bipartite Ranking using PageRank

The non-bipartite ranking technique, that we want to introduce in this section, is based on PageRank. In this section, we will finally use PageRank in order to compute a ranking that is more convincing with respect to the following condition than the average ranking from the previous section. The condition is a lesson learned in section 3.2.1:

The weight of a rating, that was made by someone who received good ratings, should be higher than the weight of a rating that was made by someone who got bad ratings.

When looking at the PageRank equation in figure 3.6, we can see, that the PageRank of a node i depends on the PageRank of the nodes that have an edge that ends in node i. In other words: The reputation of a node flows uniformly distributed over each outgoing edge to the following node(s). In the students and lecturers application, there are weighted edges (in commemoration: The weight of an edge corresponds to the given grade). However, every graph with weighted edges can be translated into a graph with uniformly-weighted edges by adding additional edges, where the weight of an edge is higher than 1 (see figure 3.7). That causes, that the higher the weight of an edge, the more reputation will flow to the following node(s).

$$PR_i = \frac{1-d}{N} + d \sum_{\forall j \in \{(j,i)\}} \frac{PR_j}{C_j} \tag{3.1}$$

Figure 3.6: PageRank equation

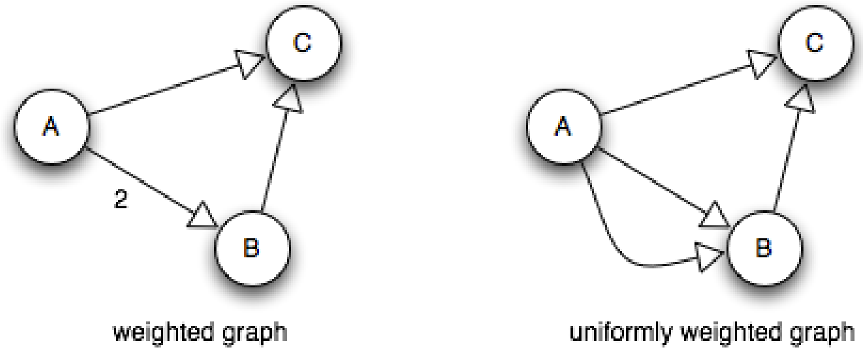

Figure 3.7: Translating a weighted graph into a uniformly weighted graph

We are going to compute the PageRank by using the power method [6]. The following vector is the PageRank solution vector of the rating table in 3.3.

$$\begin{pmatrix} 0.17864 \\ 0.16068 \\ 0.16068 \\ 0.10137 \\ 0.39863 \end{pmatrix} \tag{3.2}$$

Figure 3.8: The solution vector of the PageRank computation of the rating table of figure 3.3.

The following tables show the results of the non-bipartite PageRank ranking as well as the average ranking. Because the rankings have different scalings, we normalized them: Each result value was scaled, so that the sum of them equals 1.

Students	non-bipartite PR	normalized	average ranking	normalized
S_1	0.17864	0.3573	8.5	0.2786
S_2	0.16068	0.3213	11	0.3606
S_3	0.16068	0.3213	11	0.3606

Figure 3.9: Non-bipartite PageRank and average ranking for students

Lecturers	non-bipartite PR	normalized	average ranking	normalized
D_1	0.10137	0.2027	4	0.25
D_2	0.39863	0.7973	12	0.75

Figure 3.10: Non-bipartite PageRank and average ranking for lecturers

We now want to compare the results of the rankings. Take a look at the line charts in figure 3.11 and 3.12. The linecharts visualize the results of the average ranking and the non-bipartite PageRank ranking, in order to compare them. The line charts show the normalized values.

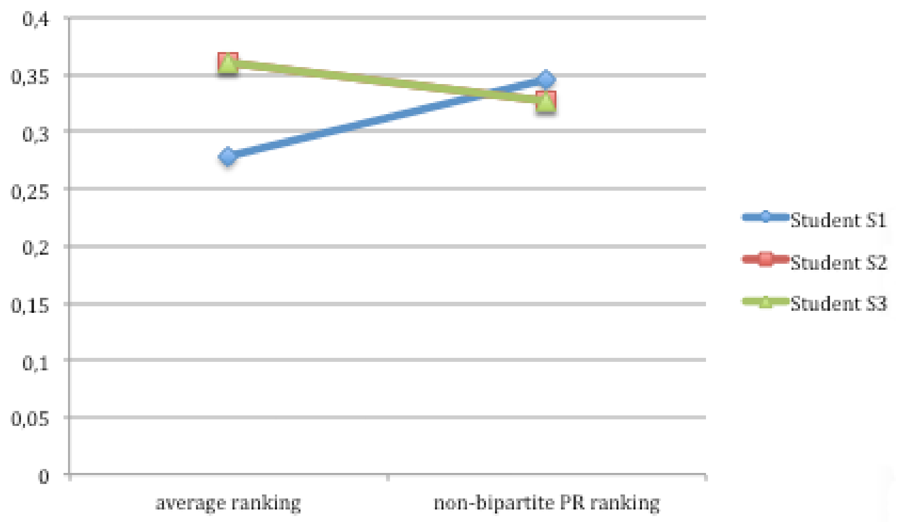

Figure 3.11: visualization of the results of the average and non-bipartite PageRank Ranking for students

While the students S_2 and S_3 got worse in the non-bipartite PageRank ranking in comparison to the average ranking, student S_1 won. The reason for that is, that S_1 got a very good rating by lecturer D_2. As already mentioned the weight of this rating depends on the ratings that D_2 received. Therefore, S_1 can win even enough to become the best rated student. S_2 and S_3 even lose in comparison to the average ranking. The reason for that is, that their highest rating (12) came from the bad lecturer. That means, that this rating's weight is lower that the weight of the lower rating (10), that was made by the better lecturer.

Let's take a look at the line chart in figure 3.12, that visualizes the results of the lecturers ranking:

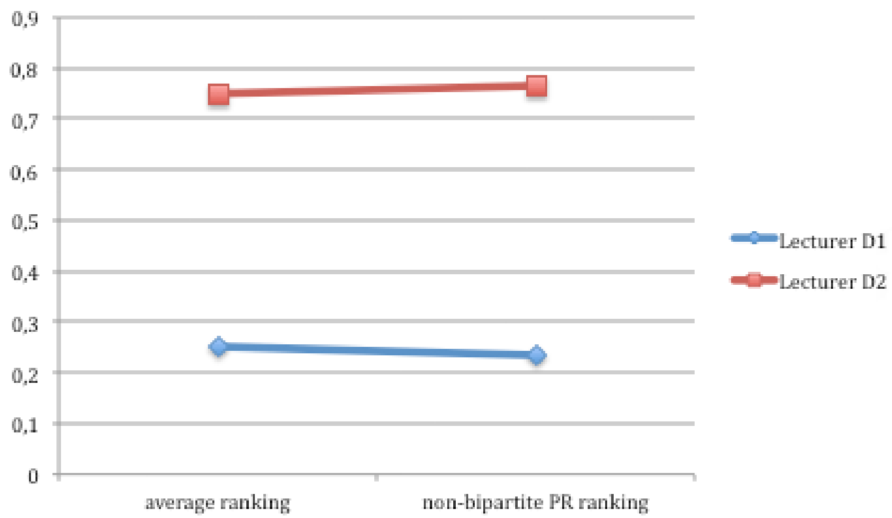

Figure 3.12: visualization of the results of the average and non-bipartite PageRank Ranking for lecturers

As we can see, lecturer D_2 won in the non-bipartite PageRank ranking in comparison to the average ranking. The reason for that is, that the good rating of the best student (S_1) has a higher weight. Lecturer D_1 however loses in the non-bipartite PageRank ranking. He lose, because he got his best ratings by the students S_2 and S_3 (12). The weight of these ratings is now lower, while the weight of the bad rating, that he got by S_1 is higher.

3.2.3 Bipartite Ranking

In the previous section, we made a ranking by using PageRank. This ranking satisfied the condition, that we formulated in section 3.2.2:

The weight of a rating, that was made by someone who received good ratings, should be higher than the weight of a rating that was made by someone who got bad ratings.

In this section, we finally want to introduce a ranking technique, that is also based on PageRank. The difference between the ranking technique from section 3.2.2 is, that this one is exclusively suitable for bipartite data.
First, we want to explain how the technique works. Then, we are going to make a ranking with the new technique and compare the results with the results of the non-bipartite PageRank ranking and the average ranking from the previous sections of this chapter.
In order to explain the bipartite approach, we want to take a look at the

matrix, that can be translated from the rating table, that we used in the previous sections.

	S_1	S_2	S_3	D_1	D_2
S_1	0	0	0	2	12
S_2	0	0	0	5	12
S_3	0	0	0	5	12
D_1	2	12	12	0	0
D_2	15	10	10	0	0

$$\longrightarrow \begin{pmatrix} 0 & 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & 5 & 12 \\ 0 & 0 & 0 & 5 & 12 \\ 2 & 12 & 12 & 0 & 0 \\ 15 & 10 & 10 & 0 & 0 \end{pmatrix}$$

Figure 3.13: Rating table and its matrix-representation

The data we work with, can always be represented as a bipartite graph. The graph is bipartite, because all participants of the ranking can be divided in two disjunct sets. And there are no ratings between members of the same group. The non-bipartite approach, that we introduced in the previous section 3.2.2 however, can be applied on every kind of graph. Therefore, we were trying to find a technique, that is exclusively suitable for bipartite graphs.

One characteristic of the matrix representation of a bipartite graph is, that such a matrix always contains two zero matrices and two non-zero submatrices. Both submatrices contain the ratings of one of the two groups. In order to have easy to remember identifiers, we define S to be the matrix, that contains the ratings, that were made by students. And L the matrix that contains the ratings, that were made by the lecturers. Figure 3.14 shows a matrix with marked submatrices S and L.

$$\begin{pmatrix} 0 & 0 & 0 & 2 & 15 \\ 0 & 0 & 0 & 5 & 15 \\ 0 & 0 & 0 & 5 & 15 \\ 2 & 12 & 12 & 0 & 0 \\ 15 & 10 & 10 & 0 & 0 \end{pmatrix} \boldsymbol{S}$$

$$\boldsymbol{L}$$

Figure 3.14: Matrix with marked submatrices S and L

The dimensions of the matrices S and L depend on the number of students and lecturers. S is a matrix with the dimension $|lecturers| \times |students|$,

while L has the symmetric dimension: $|students| \times |lecturers|$. $|students|$ denotes the amount of students that participate at the ranking, $|lecturers|$ the amount of lecturers.
We now want to define two further matrices. The matrix SL and the matrix LS. Let SL be $S \cdot L$ and $LS = L \cdot S$. Since each matrix was generated out of both submatrices, we can look separately at these two matrices, without losing the information from one of the both rating directions. SL refers to the students, while LS refers to the lecturers. For each of the matrices, we want to compute the PageRank, in order to get two separate rankings. One ranking for students by computing the PageRank for SL, and one for lecturers by computing the PageRank for LS.
We will reuse the same data, that we still used in the average ranking and the non-bipartite PageRank ranking. As already explained, the matrices SL and LS are generated by multiplying S with L, respectively L with S. You can take a look at them in the figures 3.15 and 3.16.

$$\begin{pmatrix} 229 & 174 & 174 \\ 235 & 210 & 210 \\ 235 & 210 & 210 \end{pmatrix} \tag{3.3}$$

Figure 3.15: The matrix SL

$$\begin{pmatrix} 124 & 390 \\ 130 & 525 \end{pmatrix} \tag{3.4}$$

Figure 3.16: The matrix LS

The figures 3.17 and 3.18 show the results of the non-bipartite PageRank ranking of section 3.2.2 and the results of the bipartite PageRank ranking, that we get by computing the PageRank of the SL and LS matrices. It is not necessary to normalize the results of the bipartite PageRank ranking, because the sum of the entries of the solution vector of a PageRank computation always equals 1.

Students	non-bipartite PR	normalized	bipartite PR ranking
S_1	0.17864	0.3573	0.37299
S_2	0.16068	0.3213	0.31351
S_3	0.16068	0.3213	0.31351

Figure 3.17: Non-bipartite PR and bipartite PR ranking for students

Lecturers	non-bipartite PR	normalized	bipartite PR
D_1	0.10137	0.2027	0.20737
D_2	0.39863	0.7973	0.79263

Figure 3.18: Non-bipartite PR and bipartite PR ranking for lecturers

We want to compare the results of the different rankings again. The line chart in figure 3.19 visualizes the student's results of the three rankings from this and the two previous sections.

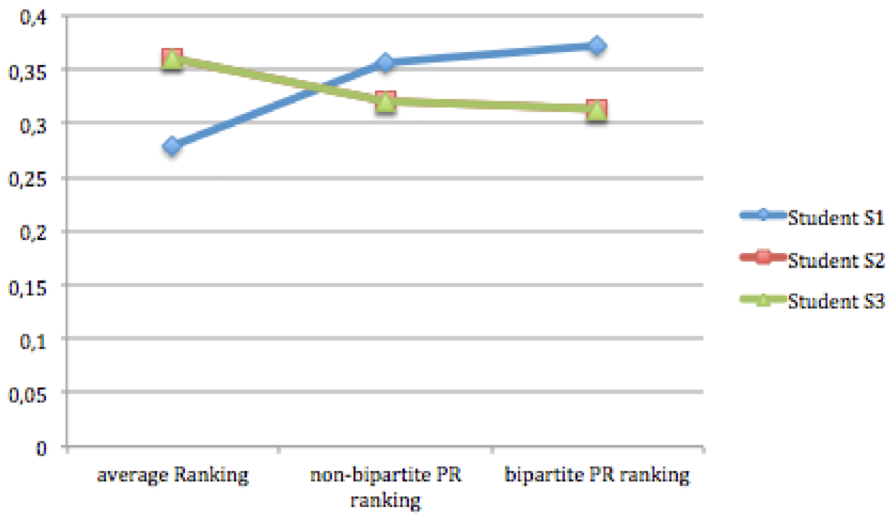

Figure 3.19: Visualization of the results of the average, non-bipartite PR and bipartite PR ranking for students

It is remarkable, that the difference between the average ranking and the bipartite ranking is larger, than the difference between average ranking and non-bipartite ranking. We made the same observation in many other scenarios. We therefore assume, that the bipartite ranking attempt emphasizes the differences between non-bipartite PageRank and average ranking.

Let us take a final look at the results of the lecturers ranking, that are visualized by the line chart in figure 3.20

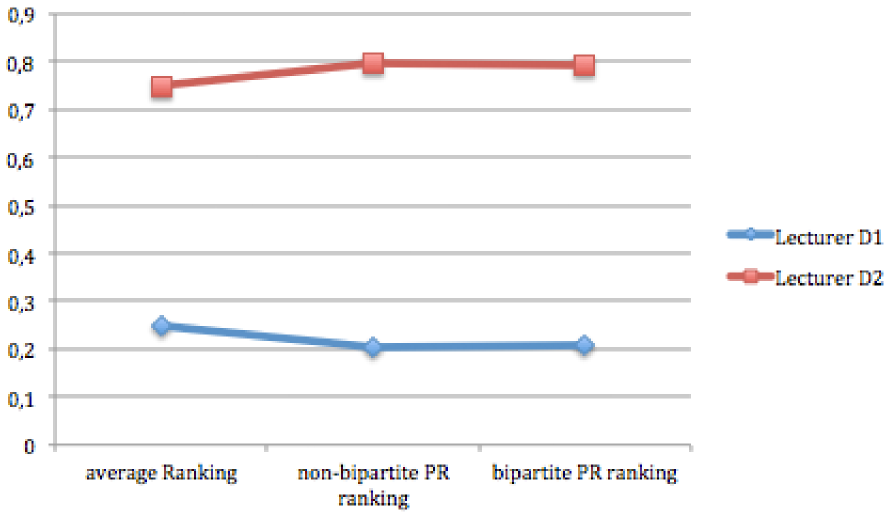

Figure 3.20: Visualization of the results of the average, non-bipartite PR and bipartite PR ranking for lecturers

The line chart shows no significant differences between the non-bipartite and the bipartite ranking. When looking at the table in figure 3.18 we see, that the differences are very small. At this time, we can not give an explanation for that, but we assume, that there are no changes under certain circumstances.We will pick the missing explanation up again in section 4.4.

Chapter 4

Validation

In this chapter, we want to take a closer look at the comparison of the non-bipartite ranking technique from section 3.2.2 and the bipartite ranking technique from section 3.2.3. More precisely, we want to verify the assumption, that the bipartite ranking emphasizes the differences between non-bipartite ranking and average ranking.
We want to do that by comparing the non-bipartite ranking technique with the bipartite ranking technique in a few hand picked scenarios. For that purpose, we developed a computer program, that can create large datasets for testing with predefined characteristics. We describe this program in section 4.1. The scenarios, their rankings and our interpretations of the results can be found in section 4.2. In the further sections of this chapter, we want to take a closer look at the damping factor and how it influences the bipartite ranking.

4.1 Validation Software

In this section, we want to introduce the validation software, that we used in order to create large datasets for the scenarios in section 4.2. Let's take a look at the rating table that we used in the previous sections:

	S_1	S_2	S_3	D_1	D_2
S_1	0	0	0	2	12
S_2	0	0	0	5	12
S_3	0	0	0	5	12
D_1	2	12	12	0	0
D_2	15	10	10	0	0

Figure 4.1: ranking table for an example containing 3 students and 2 lecturers

The scenario in the rating table contains specific specimen of students and lecturers. Every specimen's properties are defined by the ratings that it gives and receives. One could also try to find a background story for this scenario, in order to find out how the ranking technique behaves in this special case. In this example, we simply want to demonstrate how the ranking results for student S_1 change in the three different rankings. But the scenarios in section 4.2 are based on situations, that can really occur.

Since we wanted to see how the ranking technique behaves when having a few thousands of each type, it was necessary to develop a computer program, that is able to create large datasets. In our program, we can define how a certain type of student rates a certain type of lecturer and vice versa. Furthermore, we can determine how many of each type participate in the ranking. Figure 4.3 shows a screenshot of the program's GUI.
The edit fields in the screenshot, that should remember on a matrix, can be used to define how the student types rate the lecturer types. In the right column, the amount of each type can be set. After having adjusted all parameters, the program creates the matrix, that is used for the computation. The schema of the matrix is illustrated in figure 4.2.
Finally, the program writes the matrix to a specific file. This file can be loaded by the numerical computing software Octave[1]. All ranking computations are made by Octave. We developed an Octave script, that computes two rankings for a given matrix. The first ranking computes the non-bipartite attempt from section 3.2.2, the second computes the bipartite attempt from section 3.2.3. In order to visualize the ranking results in a chart or a table, the script computes the average of the result values for each type. Finally, the script yields the three vectors, that can be interpreted as the ranking results (the bipartite attempt delivers two vectors, one vector for students, the other one for lecturers), into an XML file. That XML file can for example be read out by a chart creation program.

[1]http://www.gnu.org/software/octave/

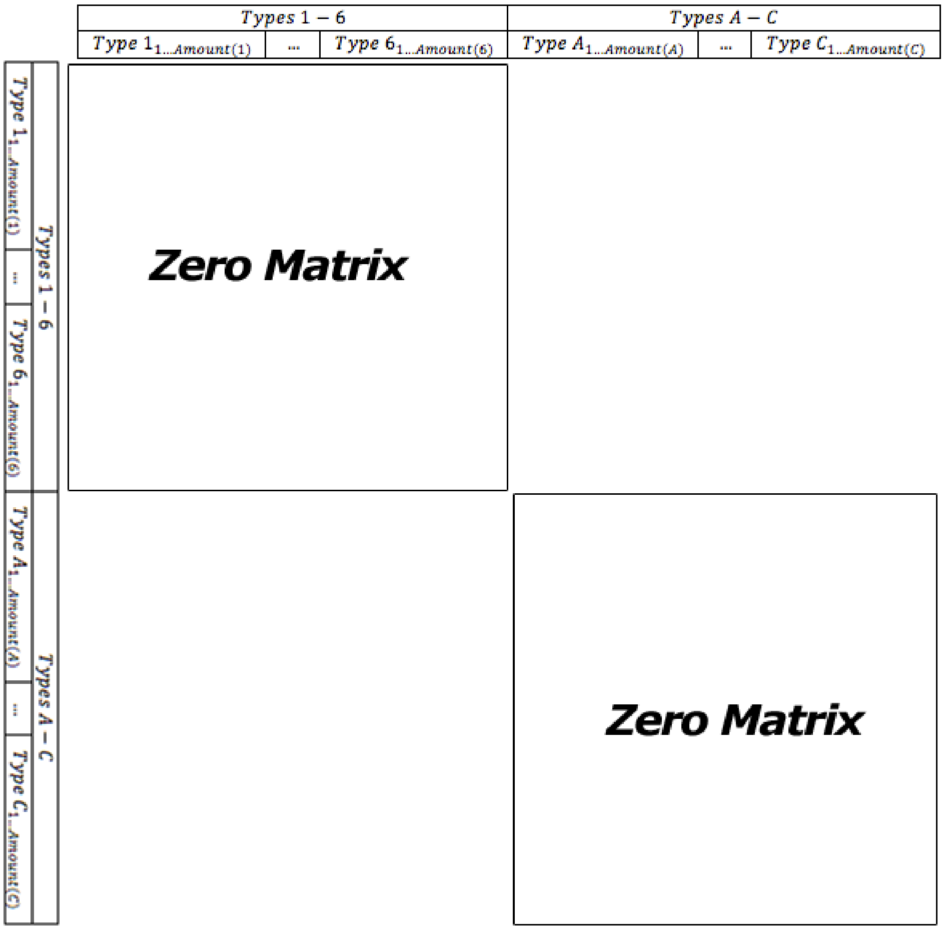

Figure 4.2: Schema of matrices, that are created by the program

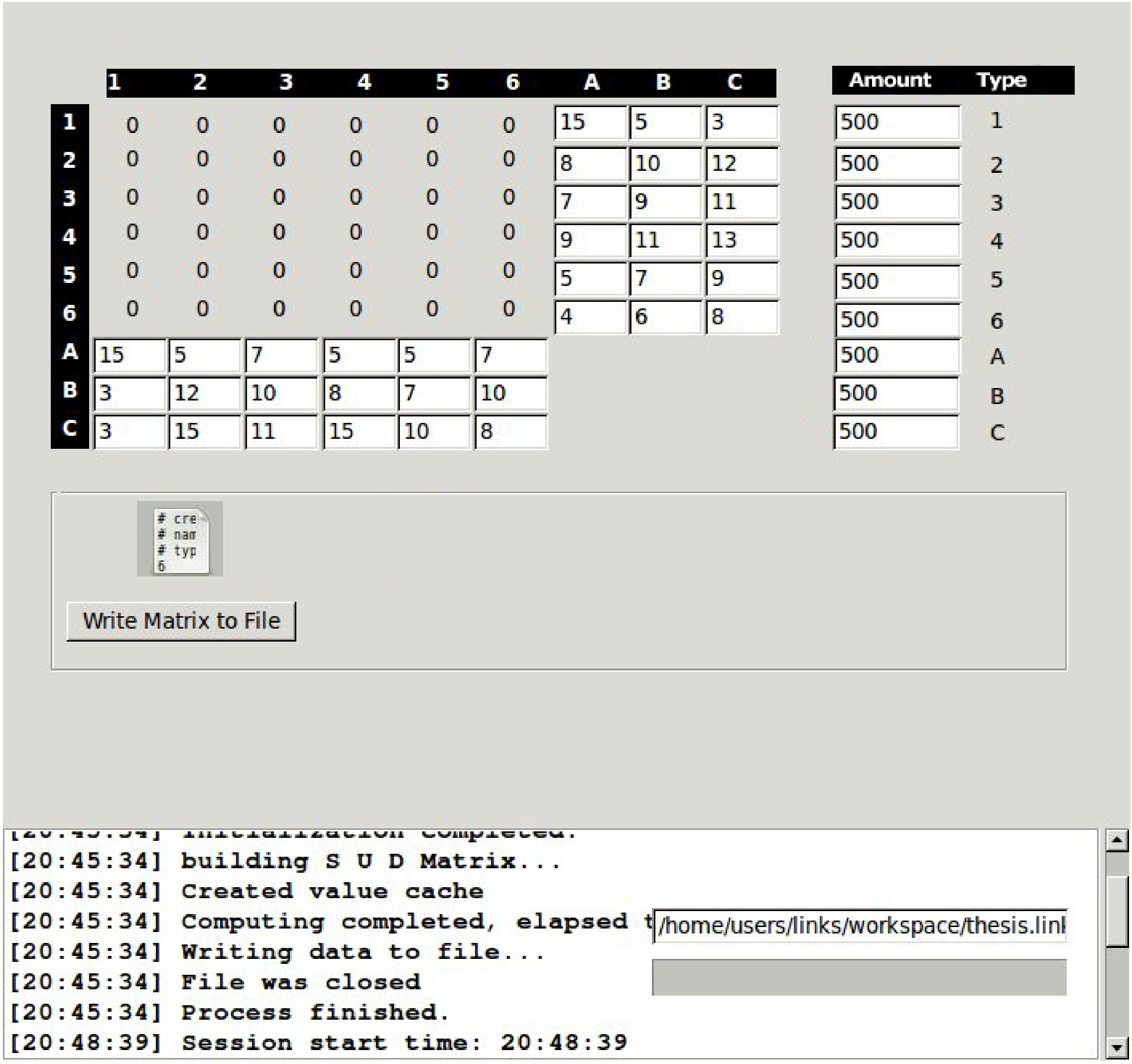

Figure 4.3: Screenshot of the program's input GUI

4.2 Scenarios

In order to verify the assumption, that the results of the bipartite attempt are stronger than the those of the non-bipartite attempt when comparing them with the results of the average ranking, we were looking at many different scenarios. For that, we took three lecturer and three students. To make it more simple, we defined three possible ratings: Good (grade 15), average (grade 10) and bad (grade 5). We used the program, that we introduced in section 4.1 of this chapter, to generate big datasets, consisting to 250 students and lecturers of each type.

In this section, we want to present two of them, that were particularly interesting, because they often occur in practice.

4.2.1 Demanding Lecturer

In this scenario, we have a very good but also demanding lecturer. The other two lecturers are average. Furthermore, there are two average students, and one student, that is very good. The important part is: The good, but demanding lecturer rates the two average students bad, because they can not satisfy his requirements. However, the good student can satisfy the requirements, and as a result, he gets the best possible rating from him. The feedback rankings of the students are symmetric: The average students are overtaxed with the demanding lecturer and as a result, they rate him bad. It is also imaginable, that they rate him bad for revenge. The good student however rates the demanding lecturer well. The scenario can be translated to the following rating table:

	S_1	S_2	S_3	D_1	D_2	D_3
S_1	0	0	0	10	10	15
S_2	0	0	0	10	10	5
S_3	0	0	0	10	10	5
D_1	15	10	10	0	0	0
D_2	15	10	10	0	0	0
D_3	15	5	5	0	0	0

Figure 4.4: ranking table for the ”demanding lecturer” scenario

We again compute the average ranking, and the two PageRank rankings, as we did it in chapter 3.2. We are particularly interested in the demanding lecturer, therefore we only discuss the lecturers ranking. The familiar chart visualizes the results:

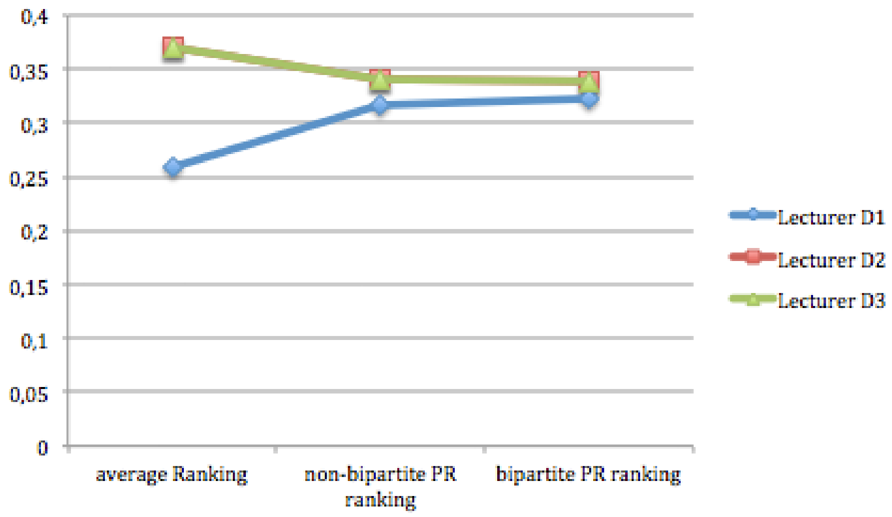

Figure 4.5: visualization of the results of the ranking results of the "Demanding lecturer"-Scenario

Of course, we strongly expected, that the demanding lecturer would improve his results in the PageRank-based rankings, because he profits by the good rating of the good student. However, we can see here again, that the improvement is stronger in our bipartite attempt, than in the non-bipartite attempt. And there is also another advantage, that becomes more obvious: In this special scenario, we were only interested in the lecturers ranking. When using the non-bipartite attempt, we have to compute the PageRank for the whole matrix, and the results includes the students ranking as well as the lecturers ranking. But in this special case, it is dispensable to compute the students ranking. When using the bipartite attempt, we can selective compute only the ranking that we need.

4.2.2 Frustrated Student

In this scenario, there are three students and three lecturer types. While the lecturers have the same level of quality, the students' performance differs. Two of them are ordinary students, one is a bit better than the other one and the ratings they get are constant. However the third student had problems with the high burden during the semester. As a result, he only visits on lecture regularly. At the end of the semester, he decides not to take part at the exams of the other two lectures. Therefore, he gets in both exams the worst possible rating (1 point). However, in the exam where he took part, he also got a bad rating (3 points). Because of that bad rating, he became frustrated. Consequently, he rates the corresponding lecturer bad as well. He did not visit the other two lectures, but the other students told him their opinions about them. So his ratings correspond to the average of the ratings of the other students. Figure 4.6 shows the corresponding ranking table.

	S_1	S_2	S_3	D_1	D_2	D_3
S_1	0	0	0	1	10	10
S_2	0	0	0	8	8	8
S_3	0	0	0	12	12	12
D_1	3	10	12	0	0	0
D_2	1	10	12	0	0	0
D_3	1	10	12	0	0	0

Figure 4.6: ranking table for the "frustrated student" scenario

Before we discuss the results of the three rankings, we want to point out, that the lecturer L_1 gets, let's say punished, because he designed the exam, where student S_1 took part. When looking at the ratings of the students S_2 and S_3 we see, that all three lecturer are actually quite homogeneous. Because of the results of the other scenarios we examined, we assume, that the punishment of S_1 does influence less in the PageRank based rankings than in the average ranking.

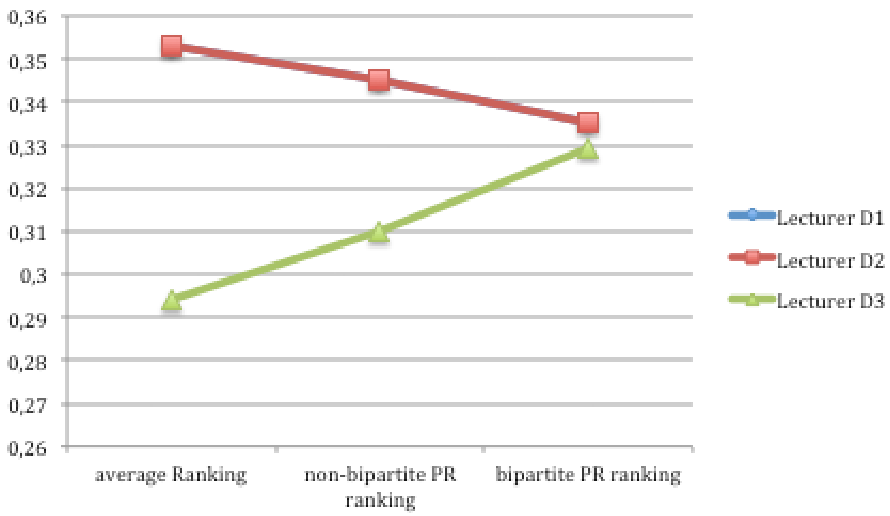

Figure 4.7: visualization of the ranking results of the "Frustrated student"-Scenario

The results confirm our assumptions. The punishment of the frustrated student became less considerable in the non-bipartite PageRank ranking, and much less in the bipartite PageRank ranking.

As well as in the previous scenario, we were able to compute the bipartite ranking only for the lecturers.

4.2.3 Homogenous Lecturers

In the last scenario, we want to check what happens when one group (in this case: the lecturers group), gets rated completely homogenous. There are again three lecturers and three students. One student gets good ratings, another one average ones and the last one bad ones. However, all three students give all lecturers the same rating. The rating table:

	S_1	S_2	S_3	D_1	D_2	D_3
S_1	0	0	0	10	10	10
S_2	0	0	0	10	10	10
S_3	0	0	0	10	10	10
D_1	14	8	5	0	0	0
D_2	12	9	4	0	0	0
D_3	15	8	3	0	0	0

Figure 4.8: Rating table of the "Homogenous lecturer" scenario

The major difference between PageRank and an average ranking is, that PageRank weights the ratings, depending on their rating values, while in a general average ranking, all ratings have the same weight. In this scenario however, the rating values of the lecturers are uniform, therefore, the non-bipartite PageRank approach should deliver the same results than the average ranking. According to our hypothesis, that the bipartite ranking emphasizes the differences between the non-bipartite and the average ranking results, we assume, that their results are equal. That is because there are no differences between the average and the non-bipartite ranking. The results proved this assumption, as you can see in the following line chart:

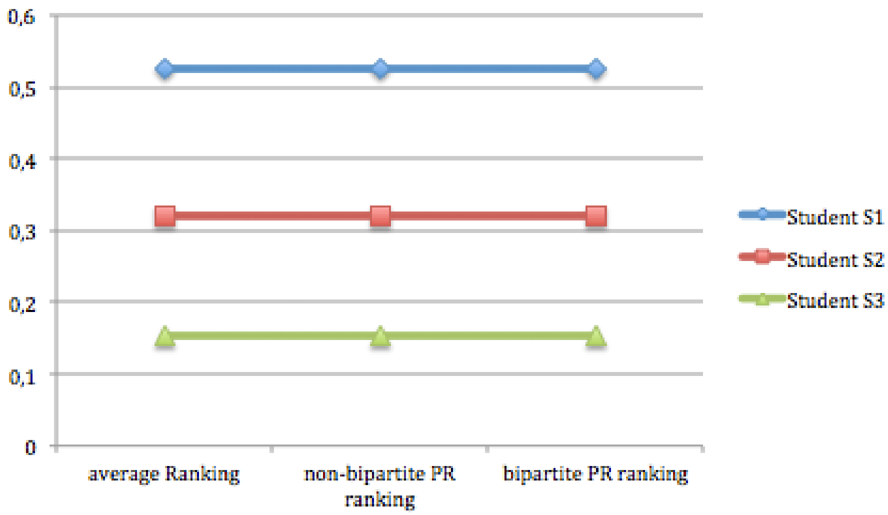

Figure 4.9: Students ranking for the "homogenous lecturers" scenario

Finally the results of the lecturers ranking: The ranking values were $\frac{1}{3}$ for each lecturer and each ranking. That is why there is only one line visible in the chart.

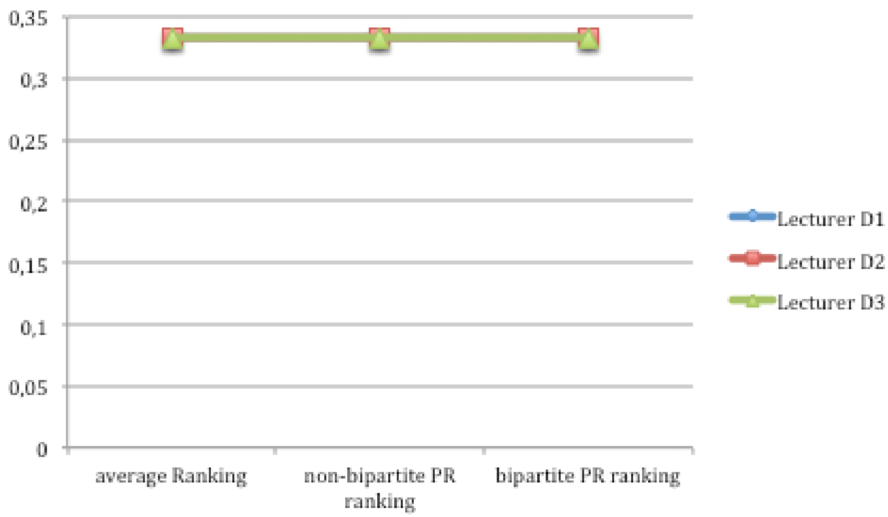

Figure 4.10: Lecturers ranking for the "homogenous lecturers" scenario

4.3 Influence of the Damping Factor

The PageRank algorithm [8] is the basis for our bipartite ranking technique. In this section we want to take a closer look at the damping factor, which is an important component of the algorithm.

The damping factor d is a free parameter ($0 < d < 1$), that controls the performance of the PageRank algorithm [5]. The underlying concept behind the damping factor is the random surfer model. In the random surfer model, the web is represented as a graph [2]. Webpages are represented as nodes. Hyperlinks between webpages edges between the corresponding nodes. By clicking hyperlinks on pages, a surfer can travel on a path trough that graph. However, a surfer can also jump to another page by typing in its URL in to the browser's address bar. The probability that a surfer jumps to a randomly chosen page can be adjusted by using the damping factor.

It is not obvious on the first sight, how the damping factor influences the results of our students and lecturers scenarios. That is because we can not translate the surfer, that travels from webpage to webpage. It is promising to take a closer look at the google matrix [7] :

$$G = dS + (1 - d)E \tag{4.1}$$

Figure 4.11: Google Matrix

In equation 4.11, d is the damping factor, S the stochastic matrix representation of the graph that shall be ranked and E the identity matrix. We can assume, that the higher the damping factor gets, the less distinguishable the results will be. In order to verify that hypothesis, we want compare the ranking results for a given scenario with $d = (0.001, 0.25, 0.5, 0.75, 0.99)$. As input for the computation, we use the rating table from the "homogenous lecturer" scenario of section 4.2.3.

	S_1	S_2	S_3	D_1	D_2	D_3
S_1	0	0	0	10	10	10
S_2	0	0	0	10	10	10
S_3	0	0	0	10	10	10
D_1	14	8	5	0	0	0
D_2	12	9	4	0	0	0
D_3	15	8	3	0	0	0

Figure 4.12: rating table of the "Homogenous lecturer" scenario

In the rating table, there are three homogenous lecturer types and three heterogeneous student types. More precisely, there is one very good student, one is average and one is bad. All lecturers get identical ratings by the students. We first look at the students' ranking results for $d = 0.001$. That is the value for d that we also used for all computations in the previous chapters.

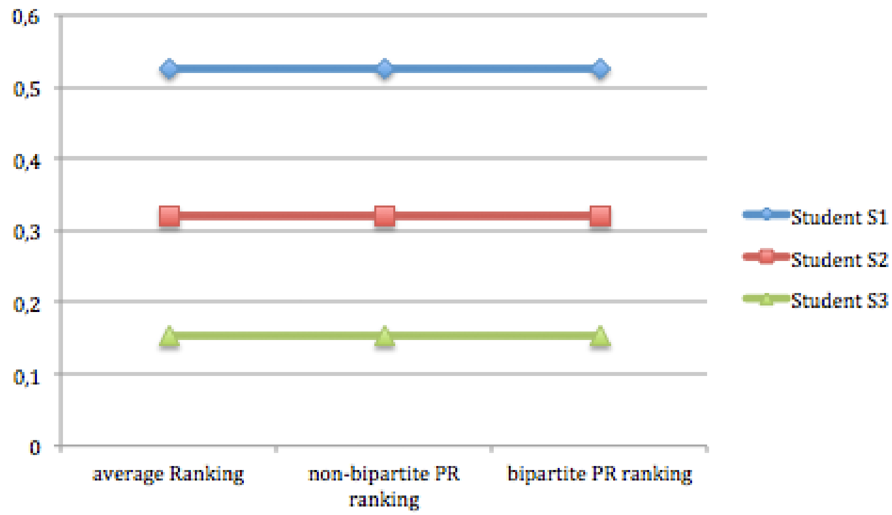

Figure 4.13: Results for students of the computations with $d = 0.001$

We can see that there are no changes differences between the results of the three rankings. The reason for that is, that differences between the average and the PageRank based rankings occur if the lecturers who rate the students are not homogenous. In this scenario they are, and therefore there are no changes.
We now want to see how the results change when increasing the damping factor. For that, take a look at the following line chart. It visualizes the results of our bipartite PageRank ranking for $d = (0.001, 0.25, 0.5, 0.75, 0.99)$.

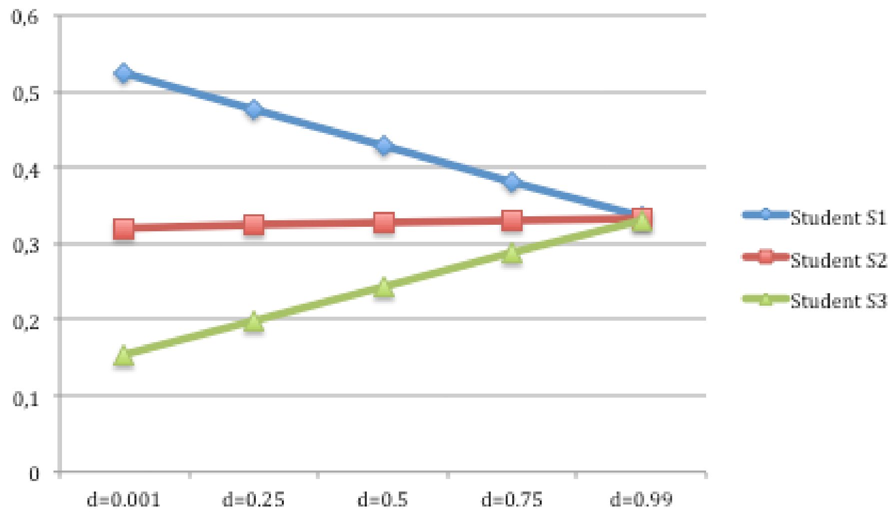

Figure 4.14: Results of the bipartite ranking with varying damping factor

The results prove our hypothesis: The higher the damping factor is, the more uniform the results are. When remembering, that the damping factor adjusted the probability that a surfer jumps randomly to any page of the graph. Translated to our students and lecturers scenarios: The damping factors causes, that edges (ratings) are randomly added to the graph. The higher the damping factor is, the more the weight of the initial ratings decreases and as a result, the results become uniform.

For the students and lecturers scenarios, we chose a very small damping factor ($d = 0.001$), because we did not want a component in the results, that is not caused by the actual ratings. In general, we recommend to choose the damping factor very small so that it doesn't influence the results.

4.4 Result Types

In section 3.2.3, we had a scenario, where the bipartite ranking did not emphasize, but equal the results of the non-bipartite ranking. And we were not able to give a satisfying explanation for that.

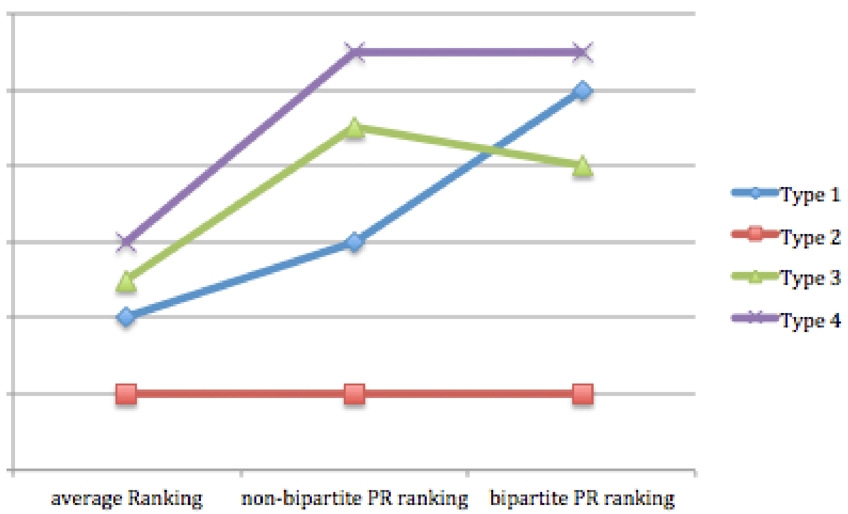

Figure 4.15: Different result types

In order to discuss the problem more generally, take a look at the line chart in figure 4.15. In the line chart, you can four result types. The types differ in the courses of their corresponding curves. We first expected, that the only existing types are 1 and 2. However, we also found one example for a type 4 result (the scenario in section 3.2.3. We were not able to find a

scenario that delivers a result of type 3, but we want to rule out its existence with certainty.
In conclusion, the questions that need to be answered are:

- Under which circumstances we get results of type 4?
- Are there scenarios, that deliver results of type 3?

In order to answer these questions, we need to find more scenarios that have the wanted result types. Therefore, we developed the concept for a program that is shortly described in the following: The program searches for type 3 and type 4 scenarios by trial and error. Adjustable parameters of the program are the number of students, number of lecturers and possible rating values (e.g.: 1...15). The program will then generate all possible scenarios for the given parameters. In the next step, the program computes the rankings and assigns them to one of the four result types. Finally, the program will deliver scenarios that most likely will show a pattern that helps answering the question under which circumstances type 3 and type 4 result types occur.

Chapter 5

Conclusion and Recommendation

We saw, that the PageRank based rankings, the non-bipartite as well as the bipartite ranking, deliver very convincing results, if there exists a reputation that flows from the one who makes a rating to the one who receives it. Ranking techniques, that don't take this reputation flow into account are inferior, in the most cases. On the other there are cases, where a reputation flow is not possible (e.g. because the rating follows strictly defined rules), not wanted (because it is defined that every rating has the same weight) or simply not possible (because of missing information about the rater's reputation).
In this thesis, we introduced a new ranking technique, that is applicable on scenarios with bipartite ratings. The non-bipartite, as well as the newly introduced bipartite ranking are applicable ranking techniques for these scenarios. If both techniques are applicable, which one should be used?
We recommend to use the bipartite ranking, due to two reasons:

- **Quality of the results**
  PageRank based rankings are used because they deliver more convincing results. More precisely: Their results differ from the results of a simple average ranking, and these differences are wanted. In the most cases (see 4.4), the bipartite ranking even emphasizes these differences.

- **Performance**
  When using the bipartite ranking, it is possible to compute only the ranking one of the two groups. There are cases, where this is very useful. When having a web application, it often happens, that a user request a specific ranking. For example, a students ranking. When using the non-bipartite ranking, it is not possible to compute the students ranking without computing the lecturers ranking. However, when using the bipartite ranking this becomes easy.

Chapter 6

Future Work

6.1 Application: Bipartite Ranking in Team Sports

In this thesis, we concentrated on students and lecturers as application for our bipartite ranking technique. But we were also looking for other domains, where our bipartite ranking is applicable. One domain, that delivers many applications is team sports. In team sports, there are usually two teams, whose members interact. It is often possible to interpret this interaction as rating. So, in the following two sections, we would like to introduce two particular interesting applications.

6.1.1 Soccer: Who is the Top Scorer?

In soccer, the question who was the best scorer of a specific tournament is rather important. Usually, answering that question is made quite simple. Each play gets one point for every scored goal. At the end of the tournament, the player with the most points is said to be the top scorer. This technique to determine the top scorer is extremely wide spread.

However, we think, that the number of scored goals does not automatically determine the best scorer. Among other things, it depends on the opponent's defence qualities how difficult it is to score a goal. That leads to a desirable condition: Not every goal should have the same weight; the weight should rather depend on the qualities of the team against it was scored. In the previous chapters of this thesis, we saw, that our bipartite ranking is able to satisfy this condition. And we would like to make a top scorer ranking of the Bundesliga season 2010/2011 that uses the bipartite ranking technique. There remain are a few considerations dealing with the two groups and the ratings:

The two groups The first group contains the scorers, more precisely: All players who scored a goal during the season. The second group contains all goalkeepers, who conceded a goal. It is appropriate to use the goalkeepers, because they represent the whole defence of a team.

The ratings When talking about the ratings, we have to distinguish the scorer to goalkeeper direction from the direction goalkeeper to scorer direction.

1. **Goalkeeper to scorer**
   The simplest possibility would be: When a specific scorer S scores a goal against a specific goalkeeper G, G will rate S with 1. The idea behind that is, that G's reputation flows to S, if S was able to beat the Goalkeeper.

   Another possibility is to define a set of rules, that determines the amount of the rating. Such a rule could be: Scoring a goal from a corner will lead to a rating of 2. Scoring by a regular shot will lead to a rating of 1.

   Another concept for such a rule could be the following: The rating depends on the distance the player has covered with the ball and number of passes before scoring. This would punish scorers who simply had to shoot the ball into the goal, and reward scorers who had no help of their teammates.

2. **Scorer to goalkeeper**
   The simple possibility would be: If a goalkeeper G clears a shoot on his goal, that was made by a scorer S, S will rate G with 1.

   Of course, it is also possible to work with rules that are similar to those we described for the goalkeeper to scorer direction.

We assume, that the rating rules will have a strong influence on the meaning of the results. Therefore they will have to be considered carefully.

6.1.2 Dodgeball: Who is the best Player?

In chapter 1, we already mentioned the game dodgeball. Dodgeball is a game played by two team, who try to eliminate the members of the opposing team by hitting them with a ball. We showed interest in dodgeball, because it is very suitable for our bipartite ranking.
Imagine a team leader wants to identify the best player in a specific match. How can he do that? Summing up the number of eliminated opponents would neglect the fact, that some players are harder to hit than others. If we assume, that a player who is good in eliminating opposing players is also hard to hit, it is very easy to use our bipartite ranking to identify the best

player of a team:

The two groups The two groups correspond to the two teams that play the match.

The ratings In contrast to the "Who is the best scorer?" ranking in section 6.1, it is not necessary to distinguish the directions. That is because we assumed, that a player who is good in eliminating opposing players is also hard to hit. That means, that we don't distinguish a player's defence behavior (avoiding to get hit) and attack behavior (trying to hit a player of the opposing team).
So we determine: When a player A hits a player B of the opposing team, there is a rating of 1 from B to A. That means, B's reputation flows to A. One can interpret that reputation as reward, that A gets for hitting B. This is valid for both rating directions.

There is one more interesting thing at dodgeball: Imagine a newspaper wants to illustrate a specific dodgeball match. A good way to do that, would be to draw a graph, that shows who eliminated who (you can such an illustration in figure 6.1). The arrows simply show how the ball was thrown (there are only shots that resulted in a hit). It is remarkable, that this illustration strongly reminds on the graph, that we would rank with our bipartite ranking. The only adaption would be to turn around the directions of the arrows.
We therefore think, that a ranking for dodgeball matches would be a good application, in order to introduce our bipartite ranking to a non-scientific audience.

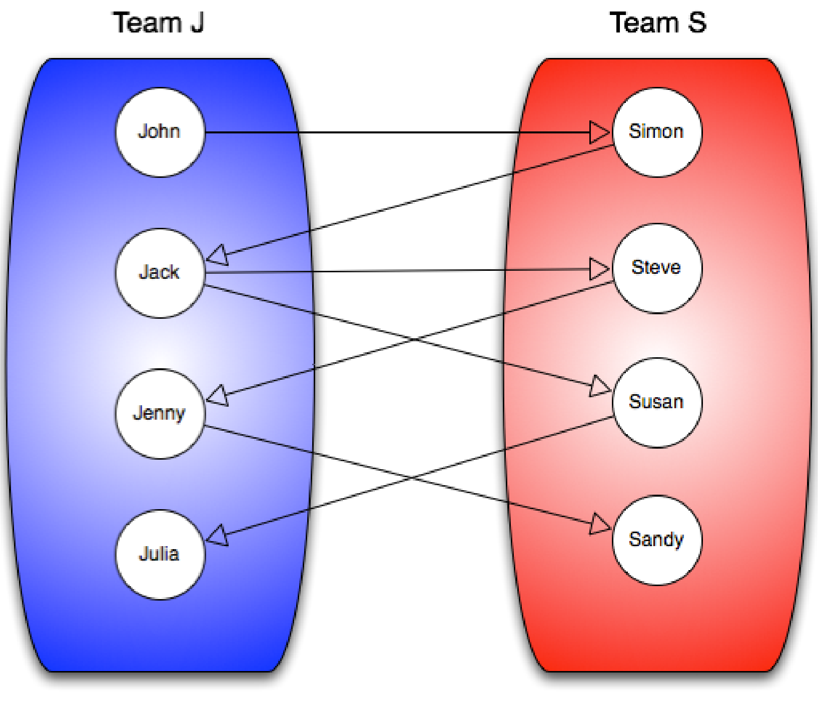

Figure 6.1: Illustration of a dodgeball match

Bibliography

[1] Alon Altman and Moshe Tennenholtz. Ranking systems: The pagerank axioms. In *In EC '05: Proceedings of the 6th ACM conference on Electronic commerce*, pages 1–8. ACM Press, 2005.

[2] L. Azzopardi, G. Kazai, S. Robertson, S. R
"uger, and M. Shokouhi. *Advances in Information Retrieval Theory: Second International Conference on the Theory of Information Retrieval, ICTIR 2009 Cambridge, UK, September 10-12, 2009 Proceedings.* Lecture Notes in Computer Science. Springer, 2009.

[3] François Bry, Norbert Eisinger, and Tim Furche. Web information systems. 2009.

[4] Andreas Hotho, Robert Jäschke, Christoph Schmitz, and Gerd Stumme. Folkrank: A ranking algorithm for folksonomies. In *Proc. FGIR 2006*, 2006.

[5] M. Lalmas, J. Jose, A. Rauber, F. Sebastiani, R. Sebastiani, and I. Frommholz. *Research and Advanced Technology for Digital Libraries: 14th European Conference, ECDL 2010, Glasgow, UK, September 6-10, 2010, Proceedings.* Lecture Notes in Computer Science. Springer, 2010.

[6] Amy Langville and Carl Meyer. Deeper inside pagerank. *Internet Mathematics*, 1(3):335–380, 01 2004.

[7] A.N. Langville and C.D. Meyer. *Google page rank and beyond.* Princeton University Press, 2006.

[8] Larry Page, Sergey Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. 1998.

[9] W. T. Tutte. *Graph theory.* Cambridge University Press, 2001.

List of Figures