Synthesis of Argumentation Graphs by Matrix Factorization

Andrea Pazienza, Stefano Ferilli, and Floriana Esposito

Dipartimento di Informatica – Università di Bari
name.surname@uniba.it

Abstract. In the phase of evaluation of accepted arguments, one may find that not all the arguments of discussion are essential when drawing conclusions. Especially when the cardinality of the set of arguments is high, the task of identifying the most relevant arguments of the whole discussion in huge Argument Systems through the analysis of its synthesis may favor better interpretability and may allow us to extract semantics that include the the strongest arguments. We propose a new matrix interpretation of argumentation graphs and exploit a matrix decomposition technique, i.e. the Singular Value Decomposition, in order to yield a synthetized argument system with only the most prominent arguments.

1   Introduction

Dung's abstract argumentation frameworks (AFs) [2] are a central concept in many argumentation formalisms and systems. Efficient and versatile methods for abstract argumentation are therefore important for further advances in the field. It is well known that AFs are usually represented as directed graphs, which play a significant role in modeling and analyzing the extension-based semantics of argumentation frameworks. Matrix representations of graphs are still in some areas the only way to represent graphs. Adjacency matrices represent adjacent vertices are capable of representing undirected and directed graphs. Matrix representations provide a bridge to linear algebra-based algorithms for graph computation. So, the potential application to argumentation frameworks is worth to expect.

The Singular Value Decomposition (SVD) [4] is the main linear technique for dimensionality reduction, aiming at producing a low-rank approximation of the original matrix with a less number of components. Our aim is therefore to exploit the matrix representation of argumentation frameworks and its low-rank approximation so as to produce a synthetized AF in order to decompose huge AFs and build a simplified ones, focusing only on arguments that are extremely useful for the evaluation process and discarding the less relevant ones, while preserving the interpretation of the whole discussion. We show that this new approach can be addressed also for generalized versions of argument graphs, such as Bipolar Argumentation Framework (BAF) [1], Weighted Argumentation Frameworks (WAFs) [3] and the recently proposed Bipolar Weighted Argumentation Frameworks (BWAFs) [5].

This paper is organized as follows. The next section addresses the problem of argument graph matrix representation and how to deal with SVD matrix factorization. A general procedure to rebuild the argument graph with the reconstructed matrix is then discussed. Section 3 shows its application to a real discussion from a Online Debating System. The last section concludes.

2 *Principal* Argument Systems

An argument graph can be represented with a slightly different version of its **adiacency matrix**. Given an Argumentation Framework $F = \langle \mathcal{A}, \mathcal{R} \rangle$, where $\mathcal{A}$ is a finite set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$, with $|\mathcal{A}| = n$, we call *Argumentation Matrix* of F the $n \times n$ matrix $M_F = [M_{ij}]$ such that for any two arguments $\alpha_i, \alpha_j \in \mathcal{A}$, the entry $M_{ij} = -1$ iff there is an attack from α_i towards α_j, otherwise the entry $M_{ij} = 0$, meaning that there is no attack relation between arguments α_i and α_j. Under this assignment, the Argumentation Matrix can be thought as a representation of the Argumentation Framework.

This representation enables to establish links between extensions (under various semantics) of the AF and to explore new properties and techniques useful to abstract argumentation puroposes. Moreover, we can extend this representation also to generalizations of argumentation frameworks.

- BAF: the entry $M_{ij} = 1$ is added to represent the support relation;
- WAF: instead of $M_{ij} = -1$, the entry of the Argumentation Matrix has a value $M_{ij} = n$, with $n \in \mathbb{R}_*^+$, that corresponds to the weight of the attack relation between two arguments;
- BWAF: entries of Argumentation Matrix have values $M_{ij} = w$, with $w \in [-1, 1]$ corresponding to the relative strength of relations between arguments.

Matrix decomposition allows us to deal with the problem of low-rank approximation. It is a minimization problem, in which the cost function measures the fit between a given matrix and an approximating matrix, subject to a constraint that the approximating matrix has reduced rank. For this purpose, we exploit SVD.

2.1 Singular Value Decomposition in a Nutshell

Let $A \in \mathbb{R}^{m \times n}$ be a matrix and $p = \min(m, n)$, the **SVD** of A is a factorization in the form $A = U \Sigma V^T$ where $U = (u_1, \ldots, u_m) \in \mathbb{R}^{m \times m}$ and $V = (v_1, \ldots, v_n) \in \mathbb{R}^{n \times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix with elements $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$. Each $\sigma_1, \ldots, \sigma_p$ is called singular value or *principal component* of A. They correspond to the square roots of A's eigenvalues.

The Eckart–Young theorem allows us to solve the problem of approximating a matrix A with another matrix $\tilde{A}$, said **truncated**, which has a specific rank k. If A is a matrix of rank $r > 1$, $A = U \Sigma V^T$ is the corresponding SVD, and $\mathcal{B} = \{B \in \mathbb{R}^{m \times n} : rank(B) \leq k\} \ \forall k \in \{1, \ldots, r-1\}$, than for all $\tilde{A} \in \mathcal{B}$, where $\tilde{A} = U_k \Sigma_k V_k^T = \sum_{i=1}^{k} a_i \sigma_i v_i^T$, it holds that $\min_{B \in \mathcal{B}} \|A - B\|_F^2 = \|A -$

$\tilde{A}\|_F^2 = \sum_{i=k+1}^r \sigma_i^2$, i.e., the k-dimensional submatrix $\tilde{A}$ of A represents the closest approximation of A between all the approximations of A with rank less than or equal to k, in Frobenius norm. In other words, the approximation is based on minimizing the Frobenius norm of the difference between A and $\tilde{A}$ under the constraint that $rank(\tilde{A}) = k$. It turns out that the solution is given by the SVD of A, namely $\tilde{A} = U_k \Sigma_k V_k^T = \sum_{i=1}^k u_i \sigma_i v_i^T$ where Σ_k is the same matrix as Σ except that it contains only the k largest singular values (the other singular values are replaced by zero).

2.2 Argument System Reconstrunction

Given the Argumentation Matrix of the argumentation graph under consideration (i.e, AF, BAF, WAF, BWAF), its low-rank approximation with the truncated SVD, reduced with only the k largest principal components, will ensure that the reconstruction will be the best approximation, and at the same time will preserve the meaning of the discussion. The problem relies on how to select the number of principal components k to keep aboard. This is tackled with the *Kaiser criterion*, which defines a rule to investigate the *scree plot* of matrix eigenvalues. It plots eigenvalues and looks for the "elbow": such a plot when read left-to-right across the abscissa can often show a clear separation in fraction of total variance where the 'most important' k components cease and the 'least important' components begin. The point of separation is often called the 'elbow'; the rest is "scree". Therefore, the number of components k that make up the "cliff" are extracted.

Now, it is possible to reconstruct the reduced approximating matrix $E \in \mathbb{R}^{n \times n}$, where n is the number of nodes of the argumentation graph, considering only its k largest principal components. Depending on the considered argumentation graph (i.e., AF, BAF, WAF, BWAF), the approximation of the corresponding reconstructed Argumentation Matrix must satisfy specific constraints, for which this matrix must be revised otherwise.

Let a, b be two arguments, then there is:

- AF: an attack relation between them iff $E_{ab} > 0.5$;
- BAF: an attack relation or a support relation between them iff, respectively $E_{ab} \leq -0.5$ or $E_{ab} \geq 0.5$, otherwise any relation is built;
- WAF: an attack relation between them iff $E_{ab} > 0$ where its weight is given by approximating the value E_{ab} to the nearest integer number;
- BWAF: an attack relation or a support relation between them with weight value equal to E_{ab} with bounds -1 or 1 iff, respectively, $-1 < E_{ab} < 0$ or $0 < E_{ab} < 1$.

The reconstructed argumentation graph will yield only the *strongest* arguments, depending on the relations survived from the operation of the SVD dimensionality reduction. In fact, we expect that this operation will delete the weakest relations from the argument graph. Then, the arguments no longer engaged in the discussion will be excluded, giving rise to a *synthetized* version of

the Argument System. We call such a graph **Principal Argument System**, which can be referred to a *Principal* AF (*P*-AF), as well as to a *Principal* BAF (*P*-BAF), *Principal* WAF (*P*-WAF), and *Principal* BWAF (*P*-BWAF).

3   Application on a Reddit Thread

In order to show the validity of such an approach in real cases, we considered a Reddit discussion of an episode of popular TV series (http://bit.ly/2fQxq4I) divided into several arguments with attacks and supports. The produced BWAF is made up of 70 arguments and 69 weighted relations, of which 52 attacks and 17 supports. The corresponding graph is reported in Figure 1.
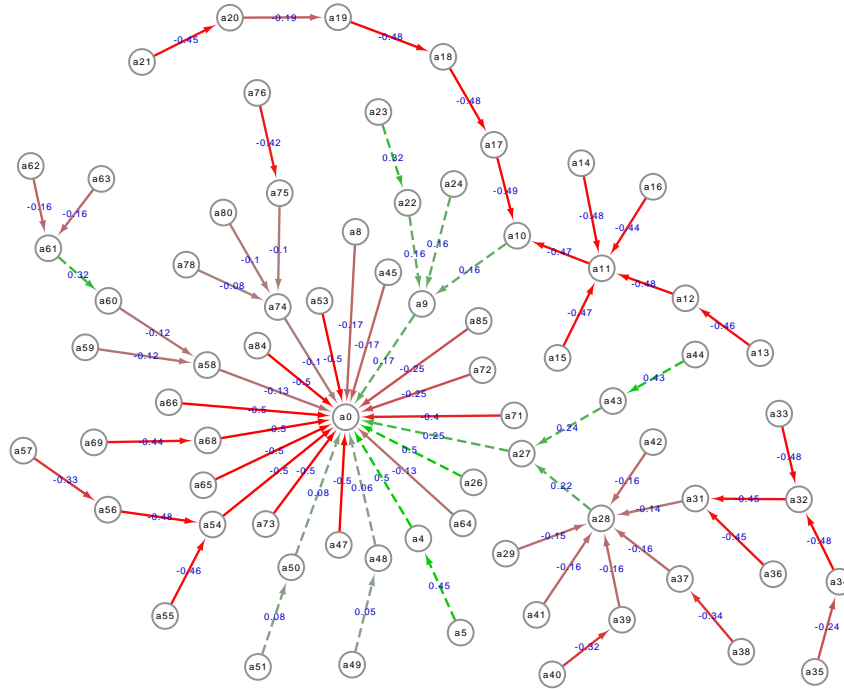

Fig. 1. BWAF representation for the considered Reddit Thread.

The synthetized version of the BWAF, i.e. the *P*-BWAF, is obtained using SVD. According to the Kaiser criterion, we reconstructed the new argument system with 25 principal components. The Principal BWAF has now 59 arguments involved in at least one relation and 58 weighted relations, of which 44 attacks and 14 supports. The corresponding graph is reported in Figure 2.

The first observation about the new synthetized *P*-BWAF is that the global meaning of the discussion has been preserved. The strongest relations continue

to exist in the revised *P*-BWAF while the weakest relations have been pruned. In particular, the following relations have been removed:

1. *support* from $a49$ towards $a48$ with strength 0.05;
2. *attack* from $a59$ towards $a58$ with strength -0.12;
3. *attack* from $a60$ towards $a58$ with strength -0.12;
4. *attack* from $a75$ towards $a74$ with strength -0.1;
5. *attack* from $a78$ towards $a74$ with strength -0.08;
6. *attack* from $a80$ towards $a74$ with strength -0.1.

From a qualitative analysis, it appears that the SVD has removed the lowest weight relations, in absolute terms, to 0.12. In particular, it has been removed about the 67% of relations with weight less than 0.12 for the starting graph. Interestingly, the SVD removed only the "peripheral" relations, i.e., those relations coming from nodes receiving no relation. This opens a new perspective about the evaluation of arguments in abstract argumentation: it turns out that sometimes, the unattacked/unsupported arguments may not be considered as winning (i.e., justified) ones, especially when the strength of their attack, or support, is extremely weak. Another interesting effect of the SVD is that even some subgraphs may have been removed. In our case, two subgraphs have been removed:
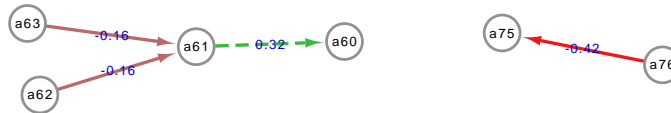
This behaviour suggests a possible strategy to determine the ideal **inconsistency budget** for WAFs and BWAFs, in order to establish a level of tolerance of the inconsistency with minimal loss of information that does not change the conclusions from those of the starting argumentation framework.

4 Conclusion

This paper propose to exploit the SVD of an argument graph to extract its syntethized version in order to highlight only the relevant arguments involved in a discussion. This approach confirms the preservation of the meaning of the whole discussion while discarding the less relevant arguments. We showed its application to a real web-based debate and discussed its effectiveness on the removed arguments and relations.

In terms of future work on this avenue of research, it would be of main interest to assess how the dimension reduction works with respect to the notion of extension-based semantics. Having introduced the basic framework of Principal Argument Systems, some important open issues arise, such as how extension-based semantics are affected by reduction, or whether arguments removed are
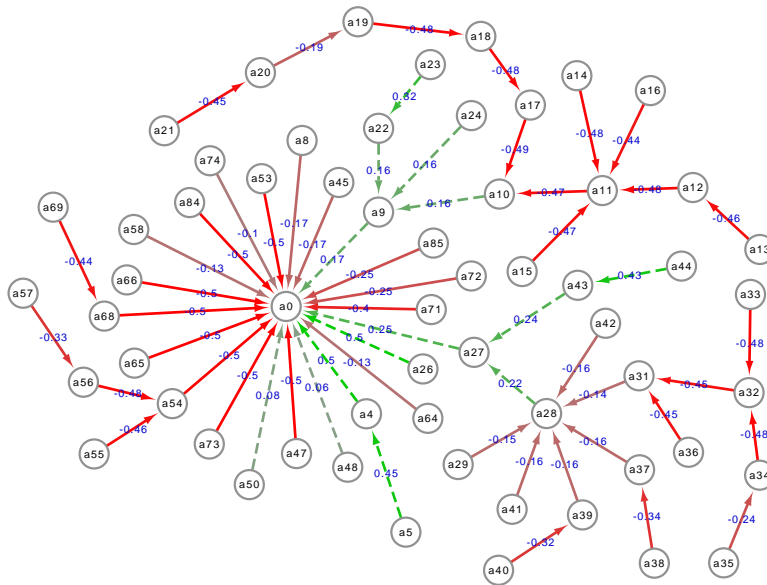
Fig. 2. *P*-BWAF representation of the resulting synthetized BWAF.

those that can be immediately flagged up as accepted/unaccepted for an extension of a given semantics. In a larger perspective, it may be useful to understand how the reduced dimension of argumentation graphs affects solvers, and, in general, whether the minimization would be helpful for the reasoning process.

Acknowledgments

References

[1] L. Amgoud, C. Cayrol, and M. Lagasquie-Schiex. On the bipolarity in argumentation frameworks. In *NMR*, pages 1–9, 2004.

[2] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence*, 77(2):321–357, 1995.

[3] P. E. Dunne et al. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence*, 175(2):457 – 486, 2011.

[4] G. H. Golub and C. Reinsch. Singular value decomposition and least squares solutions. *Numerische mathematik*, 14(5):403–420, 1970.

[5] A. Pazienza, S. Ferilli, and F. Esposito. On the gradual acceptability of arguments in bipolar weighted argumentation frameworks with degrees of trust. In *ISMIS 2017*, pages 195–204, 2017.