Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)
Doctoral Consortium Track

Sources and Information Reliability Measures*

Quentin Elsaesser
CRIL - CNRS - Université d'Artois
elsaesser@cril.fr

Abstract

More and more information is shared on the web or on social media platforms, and the information provided can be conflicting. In this case, we need to decide which information is reliable and should be taken into account. We want to define measures of reliability for each source that provides information, but also to find the truth among the conflicting information, and propose properties of these measures.

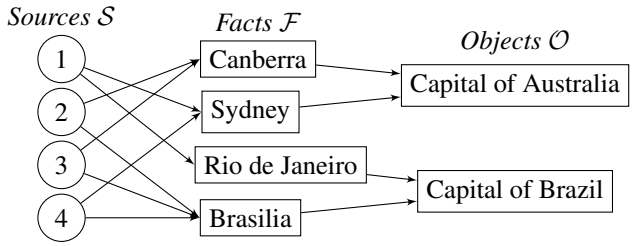

Figure 1: Sources, Facts & Objects

1 Introduction

Interaction with other agents or information sources is one of the most important ways to gain information and knowledge. When the information received is conflicting, agents who want to form opinions and resolve conflicts have the possibility to believe the most reliable sources. This is a standard way of resolving conflicts.

Truth Discovery methods aim to resolve these conflicts and find the truth among this information [Yin *et al.*, 2008; Singleton and Booth, 2022]. To achieve this task, these methods follow the idea that trustworthy sources claim believable facts. *Hubs and Authorities* [Kleinberg, 1999] ranks the web pages by defining two different types of pages. *Sums* [Pasternack and Roth, 2010] (based on Hubs and Authorities) incorporate prior knowledge and also try to find the truth.

We propose methods that allow to identify the correct answers, but also to evaluate the reliability of the sources.

2 Contribution

In [Elsaesser *et al.*, 2023], we propose a family of methods that allow to conjointly compute the reliability of a set of information sources and the confidence of the information about a set of objects, by confronting the points of view of the sources.

We keep the same structure for the data used in previous works [Yin *et al.*, 2008; Singleton and Booth, 2022] and consider three sets $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{O}$ respectively called *Sources*, *Facts* and *Objects* respectively. We define an iterative procedure to determine the reliability of the sources and find the truth among the facts (information). Our methods are characterized by the choice of a voting-scoring rule to give score to the sources depending on the ranking of the facts claimed. We wish to give an estimation of the reliability of a source, i.e. the probability of this source to find the true facts, and we use two normalization functions to ensure that the reliability of the sources is between 0 and 1. With these two normalizations, we have different information. The first one rewards sources that provide a lot of true information. A source must claim many plausible facts (i.e. facts that win the vote on its object) to receive a high reliability. The second focuses on the quality of the information. To get a high reliability, a source only needs to claim plausible facts.

We illustrate how our method works with the example in Figure 1.

We suppose that we initially have no information about the reliability of the sources or about the truth of each object. At the beginning, we assign the same reliability to all the sources. Then we compare the answers to the different questions. We rely on the idea of Condorcet's Jury Theorem arguments to find the true information and reward the sources.

For *Capital of Brazil*, the majority claims that *Brasilia* is the correct answer. *Brasilia* is the most plausible option for *Capital of Brazil* and wins the vote for its object. We reward sources 2, 3, and 4 for proposing this presumably correct answer, but there is a tie for *Capital of Australia*.

During the second iteration, the reliability of the sources claiming *Canberra* is better than the one of the sources claiming *Sydney*. Then *Canberra* becomes the most plausible option for *Capital of Australia*.

The algorithm stops when the reliability of the sources converges. In the end, *Canberra* and *Brasilia* are considered to

*Supervised by: **Patricia Everaere** and **Sébastien Konieczny**

be the truth about their respective objects. 2 and 3 are the most reliable sources. They claim the most plausible facts: *Canberra* and *Brasilia*.

3 Properties

We continue the work started in [Singleton and Booth, 2022] by proposing other properties. We propose a set of *necessary* properties that have to be satisfied by any method that aims at correctly estimating the reliability of the sources and finding the truth among the facts. We also propose a set of *optional* properties that illustrate the different behavior of some methods. Their behavior is characterized according to the scoring voting rule used or the normalization chosen.

The properties describe what the reliability of the sources should be depending on the facts they claim, and in particular whether the facts win the vote on their related object or not. We also propose properties to describe when a fact should win the vote, depending on the sources that claim it. We have a property with a special case, where the graph is composed of one object. This property is important because it states that the basic strength of a fact is given by the number of claims.

4 Experimental Study

We proceeded to an experimental evaluation of the performance of our methods for identifying the true facts and for evaluating the reliability of the sources. We evaluate our methods on generated synthetic datasets in order to perform an experimental evaluation in different settings (e.g. when the sources are reliable or not) and on two real benchmarks. We compare the results of our methods against methods from the literature Truth Finder [Yin *et al.*, 2008], Hubs and Authorities [Kleinberg, 1999], Sums [Pasternack and Roth, 2010], Unbounded-Sums [Singleton and Booth, 2022] and Voting, a method that chooses the fact with the most claims on each object. For the Truth Discovery task, we evaluate the performance of the methods with the metrics *Precision*, *Accuracy*, *Recall* and *CSI* (**C**ritical **S**uccess **I**ndex is a combination of Precision and Recall).

The real benchmarks are made for the Truth Discovery task. The benchmarks used, the *Books* and the *Flight*, are available at *http://lunadong.com/fusionDataSets.htm*. For the Truth Discovery task, our methods outperform those in the literature on the real and generated benchmarks. Our methods are more likely to find the truth when there is a tie on an object, or when the majority of sources do not claim the true fact. We also have good results for the evaluation of the reliability of the sources. We compute the averaged difference between the probability that the sources choose the true facts on each object and the reliability we obtain with our methods. Thus, this distance measures how close the estimated reliability of the sources is close to the probability. We also check the differences between the ranking of the sources obtained with the algorithm and the ranking obtained with the a posteriori probability.

For the averaged difference, the estimation of the reliability of our methods is close to the probability that the source finds the truth. For one of our methods, the reliability is even identical to the probability when the reliability of the sources is greater than 50%. This method also finds the exact ranking of the sources.

5 Conclusion and Future Works

We have introduced new methods for evaluating the reliability of the sources conjointly to the credibility of the facts in an information-based multi-agent system. In the experimental evaluations, we saw that our methods outperform methods from the literature in identifying the true facts (with the real benchmarks and with the generated datasets) and that our methods allow to correctly estimate the reliability of the sources at the same time.

They are numerous paths for future work. We want to improve the results of our methods. We can change the way we evaluate the reliability of sources, by slowly increasing or decreasing the reliability rather than completely updating the score at each iteration. For now, we assume that our objects are all on the same topic. It could also be interesting to have different topics and take into account similarities (or dependencies) between objects. With multiple topics, we can find experts in specific areas. We have no information about the sources, and when we start our algorithm, we give all sources the same reliability. The results could be different if we had a priori information about the sources. We also want to apply our methods on logical formulae instead of objects, and use our methods to define new judgment aggregation methods.

Acknowledgments

This work has benefited from the support of the AI Chair BE4musIA of the French ANR (ANR-20-CHIA-0028).

References

[Elsaesser *et al.*, 2023] Quentin Elsaesser, Patricia Everaere, and Sébastien Konieczny. S&f : Sources and facts reliability evaluation method. In *Extended Abstract. Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023)*, May 2023.

[Kleinberg, 1999] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.

[Pasternack and Roth, 2010] Jeff Pasternack and Dan Roth. Knowing what to believe (when you already know something). In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 877–885, 2010.

[Singleton and Booth, 2022] Joseph Singleton and Richard Booth. Towards an axiomatic approach to truth discovery. *Journal of Autonomous Agents and Multi-Agent Systems*, 36(2):1–49, 2022.

[Yin *et al.*, 2008] Xiaoxin Yin, Jiawei Han, and Philip S. Yu. Truth discovery with multiple conflicting information providers on the web. *IEEE Transactions on Knowledge and Data Engineering*, 20(6):796–808, 2008.