Truth Discovery: Who to Trust and What to Believe

Doctoral Consortium

Joseph Singleton
Cardiff University
singletonj1@cardiff.ac.uk

ACM Reference Format:
Joseph Singleton. 2020. Truth Discovery: Who to Trust and What to Believe. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

There is an ever-increasing amount of data available in today's world, especially on the web, social media platforms and crowdsourcing systems. Data is available in structured and unstructured formats, and from a wide range of diverse sources. The quality and reliability of such data naturally varies due to differences in the knowledge and motivation of sources. For example, some sources share inaccurate information which they falsely believe to be true, whereas others deliberately aim to misinform. This inevitably leads to conflicting information and the following question: who should we trust, and what should be believe?

Truth discovery algorithms [11, 15, 24] attempt to solve this problem by jointly estimating the trustworthiness of data sources and the likelihood a piece of information – termed a 'fact' – is true. These two measures should cohere with each other, so that a source is deemed trustworthy when it provides believable facts, and a fact is believable when it is backed up by trustworthy sources.

The problem has received increasing attention in the data mining and crowdsourcing literature recently, with many algorithms put forward. Such algorithms may use probabilistic models [14, 23], optimisation methods [13, 16] or heuristics [10, 17].

However, there are still gaps in the literature which must be addressed. In my PhD I aim to investigate such gaps from a formal point of view, using techniques and ideas from other areas of the AI and KR literature. Relevant directions for research are outlined in the next section, before a brief discussion of my work in this area so far.

2 RESEARCH DIRECTIONS

Whilst existing truth discovery algorithms perform well experimentally, there remain foundational and practical issues to be addressed. Note that the following list of issues is not exhaustive.

2.1 Theoretical Foundations

Broadly speaking, there is a lack of theoretical foundations underpinning truth discovery. Much work in the literature has a practical focus, with emphasis on construction of new algorithms and empirical evaluation of performance using large test datasets. It is true that theoretical work does exist – e.g. considerations of run-time complexity [25] and proofs of optimality in the context of a particular probabilistic framework [22] – but such results are limited in scope since they apply only to the specific model of truth discovery in question. Put simply, existing theoretical work allows one to prove results about a *specific algorithm*, but not about truth discovery as a whole.

Accordingly, there is a need for a general *unifying* framework in which existing and future algorithms can be defined. Such a framework should be neutral with respect to the methodology employed (e.g. a probabilistic framework or optimisation model), and would provide a common setting in which to compare different approaches. Various algorithms could then be evaluated with respect to their *theoretical properties*, instead of the somewhat opaque measures of performance obtained through empirical evaluation. It is hoped that this will provide stronger justification for and against particular approaches for a particular situation, and lead to insights not attainable through empirical means.

Such formal foundations are popular for related problems in *social choice theory*. Particularly relevant problems are judgement aggregation [8] and voting theory [26], wherein the *axiomatic approach* has been applied to great success. Here desirable properties (called axioms) of aggregation methods or voting mechanisms are formally stated, and the interactions between them are studied. Such analysis can lead to interesting and deep results regarding the problem as a whole; e.g. *impossibility results* which show that a combination of intuitively desirable properties are impossible to satisfy simultaneously. Taking an axiomatic social choice perspective on truth discovery has been the focus of my initial work, which is outlined later in this document.

It should be noted that the discussion in this section is not an attempt to diminish the importance of practically focussed work in truth discovery. Indeed, it is hoped that this strand of research will complement practical work and allow new algorithms to be developed with stronger theoretical backing.

2.2 Epistemic Considerations

Due to the unsupervised nature of many truth discovery algorithms, there is a risk that they simply find *consensus* amongst sources as opposed to the *truth*. Indeed, a common principle in truth discovery is that true facts are those claimed by trustworthy sources, and trustworthy sources are those who claim true facts. This mutual dependence is usually resolved by iteratively estimating source trustworthiness based on the current estimate for the true facts, and then estimating the true facts based on the updated trustworthiness scores. This means that a group of sources in agreement across a range of issues can come out as highly trustworthy, and their facts as highly likely to be true. This is perhaps a weak notion of truth, based on consistent consensus between sources.[1] Consequently, it

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

is difficult to interpret in what sense the outputs of truth discovery are 'true'. The problem therefore needs to be investigated from a formal epistemic perspective in order to establish which notion of truth is at play.

Truth-tracking methods have been proposed for this purpose in judgment aggregation, belief revision and belief fusion [3, 9, 12, 18]. Such methods model the possible states of the real world and investigate under what conditions the true facts can be fully recovered from the multi-source input. For example, in [9] the authors state a truth-tracking postulate for belief merging which requires that a merging operator uniquely identifies the true world with limiting probability 1 as the number of agents goes to infinity, under some independence and reliability assumptions on the agents. Interestingly, merging operators with similar properties may differ in whether they can track the truth; this is evidence that the problem of finding the truth is not the same as aggregating beliefs.

A key difference in the truth discovery setting is that we expect to be dealing with *unreliable sources*. The reliability assumptions of previous work therefore need to be modified; for example one could require only that a certain proportion of sources are trustworthy, or that the unreliable sources are sufficiently unreliable so as to be distinguishable from the reliable ones. Making this precise and investigating existing algorithms with respect to truth-tracking is an interesting direction for future work.

2.3 Explainability

Existing truth discovery methods suffer from a lack of *explainability*. Indeed, while an algorithm may correctly find the most plausible situation in the context of a probabilistic model or correctly minimise an objective function in an optimisation problem, this provides nothing in the way of a *reason* for why one source was deemed more trustworthy than another.

It is true that explainability is not possible or even desirable in all cases; e.g. for non-critical applications or when the number of sources is too large. On the other hand, if the results of truth discovery are to be presented to non-technical users – e.g. websites ranked by trustworthiness in search engine results or aggregation of (potentially conflicting) news stories – explainability may be necessary. Ironically, users may have little trust themselves in a black-box truth discovery system.

Argumentation theory [4] may provide a way towards explainable truth discovery. Indeed, argumentation considers arguments for and against positions, where conflicting arguments may *attack* each other, and the semantics applied to find the 'acceptable' arguments often have simple interpretations. Discussion games [5] may be particularly suitable for this purpose.

2.4 Knowledge Representation

A final gap in the truth discovery literature concerns the way in which knowledge and belief is represented. Many existing approaches take an abstract view, where the claims made by sources are either categorical values or represented by a number in $\mathbb{R}$.

However, in many cases claims from sources must first be extracted from unstructured formats such as natural language text, and the process of mapping to abstract values may discard useful information relevant for truth discovery. For example, the source may provide reasons for making the claims it does, or the claim itself could be formed from smaller claims. Methods utilising the inherent structure of such claims have been applied to the related problems of 'fake news' detection [1, 6] and argumentation mining [7]. Applying these ideas in the multi-source setting may provide an improvement over existing truth discovery methods.

[1] Note that this weaker notion of truth may still be appropriate in some domains, e.g. subjective crowdsourcing tasks where the aim is not to find any objective truth, but filter out unreliable sources with wildly different opinions.

3 WORK SO FAR

This section outlines my work so far; namely an axiomatic truth discovery framework [19, 20] and preliminary investigation into argumentation-based truth discovery [21].

3.1 An Axiomatic Approach

In accordance with the research aims of section 2.1, in [19, 20] we initiate work on an axiomatic approach to truth discovery using the tools of social choice. The contributions of this work include:

- A formal framework for truth discovery capable of modelling many existing algorithms.
- Formulation of various axioms, many adapted from the social choice and ranking systems literature [2].
- Axiomatic characterisation of a baseline 'voting' algorithm, which leads to the first impossibility result for truth discovery.
- Preliminary analysis of existing algorithms with respect to the axioms.

Future work on the framework will incorporate truth-tracking and axioms specific to truth discovery. This is hoped to more faithfully capture the character of the problem and explore the similarities and differences with other social choice problems.

3.2 Truth Discovery and Argumentation

It is clear that there is a link on some level between truth discovery and argumentation. In both cases there are conflicts between objects – conflicting 'facts' for truth discovery and attacking arguments for argumentation. The interaction between attacking arguments is well-studied in argumentation, leading to a range of semantics for finding combinations of acceptable arguments. An interesting question is whether truth discovery can be formulated as a special case of argumentation. In this case semantics would automatically induce new truth discovery algorithms. As mentioned in section 2.3, this could bring the benefit of explainability to truth discovery.

Bipolar argumentation, where arguments may support as well as attack each other, could be particularly useful. Indeed, in a truth discovery datasets there are conflicts between facts, but sources *support* facts which they believe to be true. Towards exploring this idea, my preliminary work has explored the ways in which a bipolar argumentation framework can be constructed from a truth discovery dataset [21]. In particular, I have defined a mapping from our truth discovery framework of [20] to bipolar argumentation, and applied bipolar semantics to some examples. It is promising that the induced outputs make intuitive sense in the truth discovery context. More work is required to investigate the links on a deeper level and find the most appropriate argumentation formalism.

REFERENCES

[1] H. Ahmed, I. Traore, and S. Saad. 2017. Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In *Proc. ISDDC*. 127–138.
[2] A. Altman and M. Tennenholtz. 2008. Axiomatic Foundations for Ranking Systems. *J. Artif. Int. Res.* (2008), 473–495.
[3] A. Baltag, N. Gierasimczuk, and S. Smets. 2019. Truth-Tracking by Belief Revision. *Studia Logica* (2019), 917–947.
[4] P. Baroni, D. Gabbay, M. Giacomin, and L. van der Torre (Eds.). 2018. *Handbook of Formal Argumentation*. College Publications.
[5] M. Caminada. 2017. Argumentation semantics as formal discussion. *Journal of Applied Logics* (2017), 2457–2492.
[6] N. J. Conroy, V. L. Rubin, and Y. Chen. 2015. Automatic deception detection: Methods for finding fake news. *Proc. ASIS&T* (2015), 1–4.
[7] M. Dusmanu, E. Cabrio, and S. Villata. 2017. Argument Mining on Twitter: Arguments, Facts and Sources. In *Proc. EMNLP*. 2317–2322.
[8] U. Endriss. 2016. Judgment Aggregation. In *Handbook of Computational Social Choice*.
[9] P. Everaere, S. Konieczny, P. Marquis, et al. 2010. The Epistemic View of Belief Merging: Can We Track the Truth?.. In *ECAI*. 621–626.
[10] A. Galland, S. Abiteboul, A. Marian, and P. Senellart. 2010. Corroborating Information from Disagreeing Views. In *Proc. WSDM*. 131–140.
[11] M. Gupta and J. Han. 2011. Heterogeneous Network-based Trust Analysis: A Survey. *SIGKDD Explor. Newsl.* (2011), 54–71.
[12] S. Hartmann and J. Sprenger. 2012. Judgment aggregation and the problem of tracking the truth. *Synthese* (2012), 209–221.
[13] Z. Huang, H. Ding, and J. Xu. 2019. A Faster Algorithm for Truth Discovery via Range Cover. *Algorithmica* (2019), 4118–4133.
[14] Q. Li, Y. Li, J. Gao, L. Su, B. Zhao, M. Demirbas, W. Fan, and J. Han. 2014. A Confidence-Aware Approach for Truth Discovery on Long-Tail Data. *Proc. VLDB Endow.* (2014), 425–436.
[15] Y. Li, J. Gao, C. Meng, Q. Li, L. Su, B. Zhao, W. Fan, and J. Han. 2016. A Survey on Truth Discovery. *SIGKDD Explor. Newsl.* (2016), 1–16.
[16] Y. Li, Q. Li, J. Gao, L. Su, B. Zhao, W. Fan, and J. Han. 2016. Conflicts to Harmony: A Framework for Resolving Conflicts in Heterogeneous Data by Truth Discovery. *IEEE TKDE* (2016), 1986–1999.
[17] J. Pasternack and D. Roth. 2010. Knowing What to Believe (when You Already Know Something). In *Proc. COLING*. 877–885.
[18] G. Pigozzi and S. Hartmann. 2007. Aggregation in Multiagent Systems and the Problem of Truth-Tracking. In *Proc. AAMAS*.
[19] J. Singleton and R. Booth. 2020. An Axiomatic Approach to Truth Discovery. (2020). http://users.cs.cf.ac.uk/SingletonJ1/_static/singleton_booth_axiom_full.pdf
[20] J. Singleton and R. Booth. 2020. An Axiomatic Approach to Truth Discovery (Extended Abstract). In *Proc. AAMAS*.
[21] J. Singleton and R. Booth. 2020. On the Link Between Truth Discovery and Bipolar Abstract Argumentation. (2020). http://users.cs.cf.ac.uk/SingletonJ1/_static/singleton_booth_td_arg_2020.pdf
[22] H. Xiao, J. Gao, Z. Wang, S. Wang, L. Su, and H. Liu. 2016. A Truth Discovery Approach with Theoretical Guarantee. In *Proc. SIGKDD*. 1925–1934.
[23] Y. Yang, Q. Bi, and Q. Liu. 2019. A probabilistic model for truth discovery with object correlations. *Knowledge-Based Systems* (2019), 360 – 373.
[24] X. Yin, J. Han, and P. S. Yu. 2008. Truth Discovery with Multiple Conflicting Information Providers on the Web. *IEEE TKDE* (2008), 796–808.
[25] X. Yin and W. Tan. 2011. Semi-supervised Truth Discovery. In *Proc. WWW*. 217–226.
[26] W. S. Zwicker. 2016. Introduction to the Theory of Voting. In *Handbook of Computational Social Choice*.