Truth Discovery – FTS talk

Joe Singleton

November 2019

Hello

- I'm Joe
- First year PhD
- Topic for now is *truth discovery*
- This talk is preparation for a seminar – feedback welcome
- Feel free to ask questions throughout

What is Truth Discovery?

- Lots of information is available today, from many different sources
  - The web
  - Social media platforms (Twitter, Facebook, ...)
  - Crowdsourcing systems
- People often *disagree* with what is true. Who should we trust in this case, and what should we believe?
- **Truth discovery:** find *true facts* and *trustworthy data sources* when faced with conflicting information.

This talk

- Background and context to the problem
- Existing work in this area
- My (preliminary) work:
    - How is it different?
    - What have I actually done?

This talk

- Background and context to the problem
- Existing work in this area
- My (preliminary) work:
    - How is it different?
    - What have I actually done?

Setting the scene

- Information can be collected from *data sources*
  - Websites
  - Individuals
  - Crowdsourcing participants
  - Sensors
- A piece of information relates to an *object*
  - A real-world entity or question
  - E.g. How much does the UK send to the EU per week? What will the temperature be in Cardiff tomorrow?
- Different sources can provide different 'facts' for the same object
  - Conflicting statistics
  - Much conflict over 'facts' in politics
  - Low-quality sensors
- Can result from *poor or incomplete knowledge*, or *deliberate misinformation*

Example

- fullfact.org is an ‘independent fact checking charity’

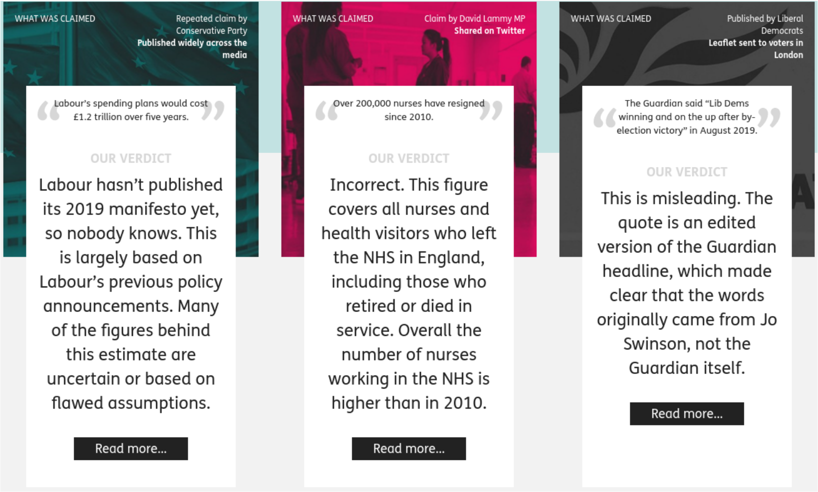


Figure 1: fullfact.org screenshot

Setting the scene (II)

- Even assuming fact-checkers are available, are they themselves to be trusted?
- Need automatic methods for finding true information
- Naive approach: take the information claimed by the most sources, i.e. perform a *vote*
- Will this work? Things to consider...
    - Large number of people today are claiming vaccines are harmful
    - A study investigated the spread of news on Twitter[1]: "Falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information"
    - Some websites copy content from each other

[1] Soroush Vosoughi, Deb Roy, and Sinan Aral. "The spread of true and false news online". In: Science 359.6380 (2018)

Setting the scene (III)

- Trouble with voting is that all sources are *equally weighted*
- It would be better to use *trust* information
    - Trustworthy sources are given more weight
    - Won't get misled by an untrustworthy majority
- Central goals of truth discovery:
    - Identify *trustworthy sources* and *believable facts*, such that trustworthy sources claim believable facts and vice versa

What does it mean to be trustworthy?

- The notion of trust is extremely important in daily life
- Trust has been studied in the social sciences, but does not have an agreed upon formal definition
- Some authors distinguish between between *trust*, *reputation* and *reliability*[2]
- Trust in daily life is often *personalised*
    - Trustworthiness is in the eye of the beholder

[2] Mohammad Momani and Subhash Challa. "Survey of trust models in different network domains". In: CoRR abs/1010.0168 (2010).

What does it mean to be trustworthy? (II)

- In contrast, truth discovery methods often seek a *global* notion of trustworthiness
- Different interpretations of trustworthiness exist in the truth discovery literature:
  - Probabilistic interpretation
  - Weights in optimisation-based methods
  - Heuristics
- Note: measures of trust are not comparable between algorithms

This talk

- Background and context to the problem
- Existing work in this area
- My (preliminary) work:
    - How is it different?
    - What have I actually done?

Existing work

- Resolving conflicts in information is not new
    - *Data fusion* considers how to combine data from multiple sources, including conflict resolution
    - *Belief revision* considers how to update existing beliefs based on new (possibly conflicting) information
- Truth discovery is distinguished by its consideration of trustworthiness
- Many algorithms proposed in recent years
    - Mostly *unsupervised*: no ground truths for objects, and no known trustworthiness values
    - Mostly *iterative*: compute trust and belief scores iteratively until convergence

Example algorithm: *Sums*

- Perhaps one of the simplest algorithms is *Sums*
- Assigns each source s a sequence of *trust scores* $\left(T_n(s)\right)_{n\in\mathbb{N}}$, and each fact f a sequence of *belief scores* $\left(T_n(f)\right)_{n\in\mathbb{N}}$.
- Initially all scores are 0.5
- Update algorithm is as follows:
    - For each source s:
        - $T_{n+1}(s) \leftarrow \sum_{f\in\text{facts}(s)} T_n(f)$
    - For each fact f:
        - $T_{n+1}(f) \leftarrow \sum_{s\in\text{src}_n(f)} T_{n+1}(s)$
    - Divide each trust and belief score by the maximum
- Repeat until convergence

This talk

- Background and context to the problem
- Existing work in this area
- My (preliminary) work:
    - How is it different?
    - What have I actually done?

Potential issues with existing work

- Lots of good algorithms, *but*...
- Many algorithms are somewhat opaque – difficult to see what the algorithm is actually *doing*
- Have to be evaluated empirically: run on a test dataset and compute accuracy
- This can make it difficult to compare algorithms:
    - Accuracy calculation depends on the dataset used
    - Algorithms may perform better or worse on different datasets
- Would be useful to have some theory behind truth discovery
    - Understand what they are doing by looking at theoretical properties
    - Make more principled comparisons
    - Deeper understanding of the problem (eventually...)

Social choice

- Theoretical analysis has been done for related problems in *social choice*
- **Social choice:** aggregate the preferences of multiple agents in a ‘fair’ way to form a social preference
    - e.g. voting: how can votes be aggregated to choose the winner of an election?
    - In our case, aggregating claims from multiple sources
- The *axiomatic approach* is popular in social choice
    - Formulate *axioms* which describe intuitively desirable properties of voting rules
    - E.g. if everyone votes for the same person, they should be elected

The axiomatic method

- Can evaluate and compare algorithms by checking which axioms are satisfied
- Common goals are *impossibility results* and *characterisation results*
- E.g. voting has *Arrow's Impossibility Theorem*
    - Three seemingly good axioms cannot hold at the same time
    - Highlights fundamental problem with voting
- E.g. Altman and Tennenholtz [3] characterised PageRank from Google
    - A ranking system coincides with PageRank iff it satisfies these axioms...
- Idea: can we give truth discovery an axiomatic treatment?

---

[3] Alon Altman and Moshe Tennenholtz. 2005. Ranking systems: the PageRank axioms.

This talk

- Background and context to the problem
- Existing work in this area
- My (preliminary) work:
    - How is it different?
    - **What have I actually done?**

What have I been doing?

- Applying axiomatic approach to truth discovery
- Defined formal framework
- Formulated some axioms
- Had a look at some existing algorithms against my axioms

The framework: what is the input to the truth discovery?

- I consider a *very basic* form of truth discovery
- We have a finite set of sources $\mathcal{S}$, facts $\mathcal{F}$ and objects $\mathcal{O}$
- Input to the problem (the dataset) is called a *truth discovery network*, and is defined as a graph

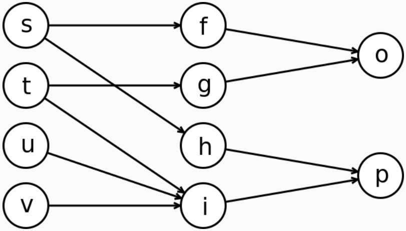


Figure 2: Example network

- Representing input as a graph is already common in the literature

Truth discovery network definition

From the paper…

Definition
A *truth discovery network* is a directed graph $N = (V, E)$ where $V = \mathcal{S} \cup \mathcal{F} \cup \mathcal{O}$, and $E \subseteq (\mathcal{S} \times \mathcal{F}) \cup (\mathcal{F} \times \mathcal{O})$ has the following properties:

1. For each $f \in \mathcal{F}$ there is a unique $o \in \mathcal{O}$ with $(f, o) \in E$, denoted $\text{obj}_N(f)$. That is, each fact is associated with exactly one object.
2. For $s \in \mathcal{S}$ and $o \in \mathcal{O}$, there is at most one directed path from s to o. That is, sources cannot claim multiple facts for a single object.
3. $(\mathcal{S} \times \mathcal{F}) \cap E$ is non-empty. That is, at least one claim is made.

We will say that s *claims* f when $(s, f) \in E$. Let $\mathcal{N}$ denote the set of all TD networks.

The framework: what is the output?

- Most algorithms give output as numeric *trust scores* and *belief scores*
- Since scores are not comparable across algorithms, we are only concerned with the *ranking* that is induced by the scores
- Output should be therefore be a pair of rankings:
  - Source ranking tells us who is *more trustworthy*
  - Fact ranking tells us which fact is *more believable*
- In the previous example, *Sums* gives the rankings

$$s < u = v < t$$

$$f = h < g < i$$

- Algorithms are represented in the framework as functions, and are called *truth discovery operators*

From the paper…

Notation
For a set X, let $\mathcal{L}(X)$ denote the set of all total preorders on X, i.e. the set of transitive, reflexive and complete binary relations on X.

Definition
A *truth discovery operator* T is a mapping $T : \mathcal{N} \to \mathcal{L}(\mathcal{S}) \times \mathcal{L}(\mathcal{F})$. We shall write $T(N) = (\sqsubseteq_N^T, \preceq_N^T)$, i.e. $\sqsubseteq_N^T$ is a total preorder on $\mathcal{S}$ and $\preceq_N^T$ is a total preorder on $\mathcal{F}$.

Axioms

- The framework provides the definitions required to formally state axioms
- Most axioms adapted from social choice
- I will only mention the important ones...

Coherence

- Recall that axioms are supposed to represent intuitive *desirable properties* of operators
- A key principle of truth discovery is that trustworthy sources make believable claims, and vice versa
- The trust and belief rankings need to *cohere* in this sense
- This idea is hard to pin down in general, but we can do so in specific cases...

Coherence (II)

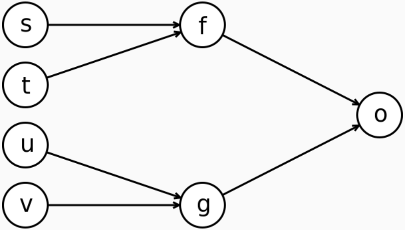


Figure 3: Coherence motivating example

- This idea comes from axiomatic analysis of ranking systems under the name *transitivity* [4]
- We consider this the most important axiom

[4] Alon Altman and Moshe Tennenholtz. 2008. Axiomatic Foundations for Ranking Systems

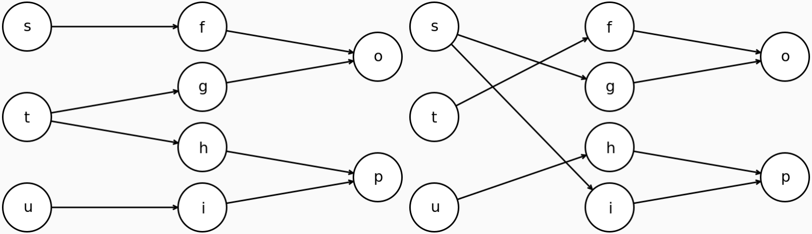
Definition
Let T be a TD operator, N be a TD network and $Y, Y' \subseteq \mathcal{F}$. We shall say Y is *less believable* than Y' with respect to N and T if there is a bijection $\phi : Y \to Y'$ such that $f \preceq_N^T \phi(f)$ for each $f \in Y$, and $\hat{f} \prec_N^T \phi(\hat{f})$ for some $\hat{f} \in Y$.

For $X, X' \subseteq \mathcal{S}$ we define X *less trustworthy* than X' with respect to N and T in a similar way.

Axiom (Coherence)
For any network N, $\text{facts}_N(s_1)$ less believable than $\text{facts}_N(s_2)$ implies $s_1 \sqsubset_N^T s_2$, and $\text{src}_N(f_1)$ less trustworthy than $\text{src}_N(f_2)$ implies $f_1 \prec_N^T f_2$.

Symmetry

- Rankings should depend on the *structure* of the network, not the *names* of sources and facts
- Consider swapping s with t and h with i:



- The structure is the same in each case, just different labels
- We should have $s \sqsubseteq_N t$ iff $t \sqsubseteq_{N'} s$ and $h \preceq_N i$ iff $i \preceq_{N'} h$
- Prevents operators being biased towards or against particular sources

Definition

Two TD networks N and N' are *equivalent* if there is a graph isomorphism π between them that preserves sources, facts and objects, i.e., $\pi(s) \in \mathcal{S}$, $\pi(f) \in \mathcal{F}$ and $\pi(o) \in \mathcal{O}$ for all $s \in \mathcal{S}$, $f \in \mathcal{F}$ and $o \in \mathcal{O}$. In such case we write $\pi(N)$ for N'.

Axiom (Symmetry)

Let N and $N' = \pi(N)$ be equivalent networks. Then for all $s_1, s_2 \in \mathcal{S}$, $f_1, f_2 \in \mathcal{F}$, we have $s_1 \sqsubseteq_N^T s_2$ iff $\pi(s_1) \sqsubseteq_{N'}^T \pi(s_2)$ and $f_1 \preceq_N^T f_2$ iff $\pi(f_1) \preceq_{N'}^T \pi(f_2)$.

- We don't want *Voting*, but more support is better in some sense...
- If f is at least as believable as g and extra support for f comes in, f becomes *strictly* more believable

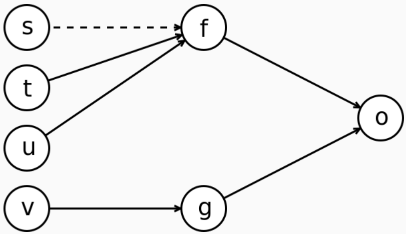


Figure 4: Monotonicity motivating example

Axiom (Monotonicity)
Suppose $N \in \mathcal{N}$, $s \in \mathcal{S}$, $f \in \mathcal{F} \setminus \mathrm{facts}_N(s)$. Write E for the set of edges in N, and let N' be the network in which s claims f; i.e. the network with edge set

$$E' = \{(s, f)\} \cup E \setminus \{(s, g) : g \neq f, \mathrm{obj}_N(g) = \mathrm{obj}_N(f)\}$$

Then for all $g \neq f$, $g \preceq_N^T f$ implies $g \prec_{N'}^T f$.

Independence

- Notion of *independence* is important: the ranking of a source/fact should only depend on the stuff that is relevant to it

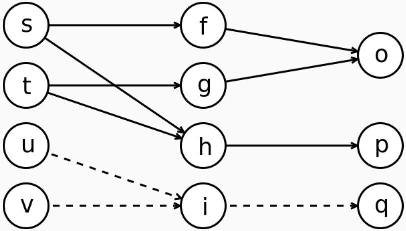
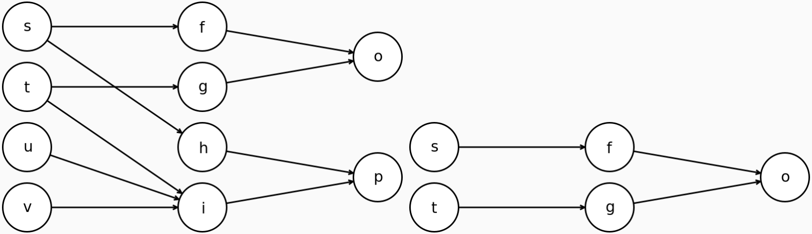


Figure 5: Independence motivating example

- e.g. are *u* and *v* relevant to *s*?

Per-object Independence (POI)

- First stab at independence, obtained by translating social choice (esp. voting) versions of independence
- If facts and sources for object o are the same in N and N', the ranking of o's facts is the same



Axiom

Let $o \in \mathcal{O}$ and write $\mathrm{obj}_N^{-1}(o) \subseteq \mathcal{F}$ for the set of facts for o in a network N. Suppose N_1, N_2 are networks such that $F_o = \mathrm{obj}_{N_1}^{-1}(o) = \mathrm{obj}_{N_2}^{-1}(o)$ and $\mathrm{src}_{N_1}(f) = \mathrm{src}_{N_2}(f)$ for each $f \in F_o$. Then the restrictions of $\preceq_{N_1}^T$ and $\preceq_{N_2}^T$ to F_o are equal; that is, $f_1 \preceq_{N_1}^T f_2$ iff $f_1 \preceq_{N_2}^T f_2$ for all $f_1, f_2 \in F_o$.

Is POI a good idea?

- POI means we cannot use inter-object links
- With Symmetry and Monotonicity, this is very bad: it implies *Voting* behaviour within the facts for each object

Theorem
Let T be any operator satisfying Symmetry, Monotonicity and POI.

Then for any network N, object o and facts f, g *for o, we have*

$$f \preceq_N^T g \text{ iff } |\mathrm{src}_N(f)| \leq |\mathrm{src}_N(g)|$$

- Note: It is possible to strengthen POI – to what I call *Strong Independence* – to get *Voting*-like behaviour for *any* two facts: we have found an axiomatic characterisation of *Voting*

Impossibility

- Remember Coherence is our key axiom, which *Voting* fails
- Symmetry, Monotonicity and POI imply *Voting*-like behaviour
- Symmetry, Monotonicity, POI and Coherence? No

Theorem
There is no operator satisfying Coherence, Symmetry, Monotonicity and POI.

Impossibility (II)

- Counterexample is shown below

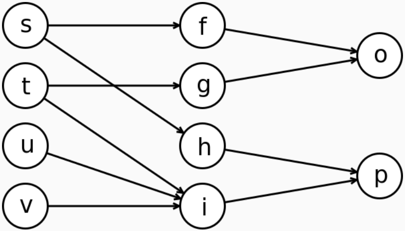


Figure 6: Counterexample used in the proof

- As far as we know, this is the first impossibility result for truth discovery

- POI is *not* desirable since it rules out using indirect links
- Our final version of independence is very weak: two nodes are relevant to each other if there is a path between them, i.e. if they are in the same *connected component* of the graph

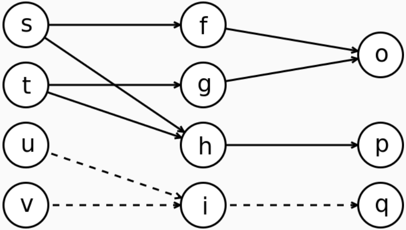


Figure 7: Independence example

Axiom (Independence)
For any TD networks N_1, N_2 with a common connected component G, the restrictions of $\sqsubseteq^T_{N_1}$ and $\sqsubseteq^T_{N_2}$ to $G \cap \mathcal{S}$ are equal, and the restrictions of $\preceq^T_{N_1}$ and $\preceq^T_{N_2}$ to $G \cap \mathcal{F}$ are equal; that is, $s_1 \sqsubseteq^T_{N_1} s_2$ iff $s_1 \sqsubseteq^T_{N_2} s_2$ and $f_1 \preceq^T_{N_1} f_2$ iff $f_1 \preceq^T_{N_2} f_2$ for $s_1, s_2 \in G \cap \mathcal{S}$ and $f_1, f_2 \in G \cap \mathcal{F}$.

Satisfaction of the axioms

- Those are the important axioms. Are they satisfied by actual truth discovery algorithms?

	Voting	SC-Voting	Sums	U-Sums
Coherence	X	X	✓	✓
Symmetry	✓	✓	✓	✓
Mon.	✓	✓	X	?
POI	✓	✓	X	X
Indep.	✓	X	X	✓

Table 1: Satisfaction of the axioms for the various operators

That’s all!

- Thanks for listening
- Questions?