A Probabilistic Model with Commonsense Constraints for Pattern-based Temporal Fact Extraction

Yang Zhou[1]**, Tong Zhao, Meng Jiang**
Department of Computer Science and Engineering
University of Notre Dame, Notre Dame, IN 46556
{yzhou24, tzhao2, mjiang2}@nd.edu

Abstract

Textual patterns (e.g., `Country`'s president `Person`) are specified and/or generated for extracting factual information from unstructured data. Pattern-based information extraction methods have been recognized for their efficiency and transferability. However, not every pattern is reliable: A major challenge is to derive the most complete and accurate facts from diverse and sometimes conflicting extractions. In this work, we propose a probabilistic graphical model which formulates fact extraction in a generative process. It automatically infers true facts and pattern reliability without any supervision. It has two novel designs specially for temporal facts: (1) it models pattern reliability on two types of time signals, including temporal tag in text and text generation time; (2) it models commonsense constraints as observable variables. Experimental results demonstrate that our model significantly outperforms existing methods on extracting true temporal facts from news data.

1 Introduction

Temporal fact extraction is to extract (entity, value, time)-factual tuples from text data (e.g., news, tweets) for specific attributes. It acts as one of the fundamental tasks in knowledge base construction, knowledge graph population, and question answering. For example, if we were interested in *country's president*, the entity would be of type `Location.Country`, the value would be of type `Person`, and the time would be a valid year in the person's presidential term. Thanks to name entity recognition (NER) and typing systems (Del Corro et al., 2015), pattern-based information extraction methods generate patterns consisted of entity types (Jiang et al., 2017; Li et al., 2018;

Reimers et al., 2016). They are widely used for good transferability across domains and datasets, unsupervised manner requiring no or very few annotations, and high efficiency. The typed patterns give only the association between entity and value. Two types of time signals can be attached to the pairs, forming temporal triples: One is temporal tag in text, e.g., the year tag next to the entity/value mentions in the sentence; the other is text generation time, i.e., the year the text document was posted. For example, given two sentences:

1) "... *The former* <u>French</u> [`Country: France`] *president* <u>Jacques Chirac</u> [`Person`]*, a self-styled affable rogue who was head of state from* <u>1995</u> *[temporal tag] to 2007 ...*" (posted on Sept. 26, <u>2019</u> [text generation time])

2) "... <u>*Emmanuel Macron*</u> [`Person`]*, now President of* <u>France</u> [`Country`]*, graduated from ENA in* <u>2004</u> *[temporal tag] ...*" (posted on Sept. 19, <u>2019</u> [text generation time])

Pattern-based methods discover two patterns:

- P1: former `Country` president `Person`
- P2: `Person`, now president of `Country`,

Then the methods can extract the following tuples. We label ✔ and ✗ for correct tuples and incorrect ones, respectively:

✔ (France, Jacques Chirac, 1995): P1 and temporal tag;

✗ (France, Jacques Chirac, 2019): P1 and text gen. time;

✗ (France, Emmanuel Macron, 2004): P2 and temporal tag;

✔ (France, Emmanuel Macron, 2019): P2 and text gen. time.

We have the following observations:

- **O1:** Not every pattern is reliable: the pattern "`Person` visited `Country`" is very likely to be unreliable. Not every pattern is unreliable: the pattern "current `Country`'s president `Person`" is very likely to be reliable. The above two pattern examples are somehow half and half. So, patterns have reliability.

- **O2:** For temporal fact extraction, different types of time signals might be either reliable

or unreliable depending on the pattern. So, there is a dependency between pattern and type of time signal, in terms of reliability.

Existing truth finding approaches assumed that a structured "source-object-claim" database was given and then estimated the reliability of source for inferring whether the claim was true or false (Yin et al., 2008; Zhao and Han, 2012; Zhao et al., 2012). For example, a source could be a book seller, an object could be a book's author list, and a claim could be an author list that a seller gave for a book. One conclusion was that *probabilistic graphical models* (PGM) (Zhao and Han, 2012; Zhao et al., 2012) have advantages of estimating source reliability over the general data distributions, compared with bootstrapping algorithms (Yin et al., 2008; Li et al., 2018; Wang et al., 2019). However, PGM-based truth finding models have not yet been developed for the task of *information extraction*. Estimating the reliability of textual patterns is new (O1). Moreover, when we focus on temporal fact extraction, modeling the dependency between pattern and type of time signals is also new (O2).

In truth finding, it is critical to define conflicts. For the book seller's example, we assume that one book can have only one true author list; so if we knew one list was true, then any different list of the same book would be false. This originated from our commonsense. Fortunately, we have quite a few commonsense rules for temporal facts, i.e., specific attributes. On *country's president*, we know that

- one president serves only one country;
- one country has only one president at a time;
- however, one country can have multiple presidents in the history (e.g., USA, France).

For the attribute *sports team's player*, we have commonsense rules:

- one player serves only one club at a time;
- however, one club has multiple players and one player can serve multiple clubs in his/her career.

We generalize possible commonsense rules:

- C1: one value matches with only one entity;
- C2: one entity matches with only one value;
- C3: one value matches with only one entity at a time;
- C4: one entity matches with only one value at a time.

So, we know that the attribute *country's president* follows C1 and C4; and the attribute *sports team's player* follows C3. The third challenge (O3) is the necessity of modeling the commonsense (e.g., C1–C4) for identifying conflicts, estimating pattern reliability, and finding true temporal facts.

To address the three challenges (O1–O3), we propose a novel Probabilistic Graphical Model with Commonsense Constraints (PGMCC), for finding true temporal facts from the results from pattern-based methods. The given input is the observed frequency of tuples extracted by a particular pattern and attached with a particular type of time signal. We model information source as a pair of pattern and type of time signal. We represent the source reliability as an unobserved variable. It becomes a generative process. We first generate a source. Next we generate a (entity, value, time)-tuple. Then we generate the frequency based on the source reliability and the tuple's trustworthiness (i.e., probability of being a truth). Moreover, we generate variables according to the commonsense rules if needed – the variable counts the values/entities that can be matched to one entity/value with or without a time constraint (at one time) from the set of *true* tuples. Given a huge number of patterns (i.e., 57,472) and tuples (i.e., 116,631) in our experiments, our proposed unsupervised learning model PGMCC can effectively estimate pattern reliability and find true temporal facts.

Our main contributions are:

- We introduce the idea of PGM-based truth finding to the task of pattern-based temporal fact extraction.
- We propose a new unsupervised probabilistic model with observed constraints to model the reliability of textual patterns, the trustworthiness of temporal tuples, and the commonsense rules for certain types of facts.
- Experimental results show that our model can improve AUC and F1 by more than 7% over the state-of-the-art.

The rest of this paper is organized as follows. Section 2 introduces the terminology and defines the problem. Section 3 presents an overview as well as details of the proposed model. Experimental results can be found in Section 4. Section 5 surveys the literature. Section 7 concludes the paper.

2 Terminology and Problem Definition

2.1 Terminology

Definition 1 (Temporal fact: (entity, value, time)-tuple)**.** *Let* $\mathcal{F} = \{f_1, f_2, f_3, \dots\}$ *be the set of temporal facts. Each fact* f *is in the format of*

(entity, value, time). $\mathcal{F}$ *was extracted by textual pattern-based methods.*

Definition 2 (Pattern s). *Let* $\mathcal{P}^{(*)} = \{p_1^{(*)}, \dots\}$ *be the set of pattern source, here* $* \in \{post, tag\}$ *stands for the type of time signal (i.e., "text gen. time" and "temporal tag"). One pattern paired with different types of time signals will be treated as different pattern sources.*

Definition 3 (Extraction). *Let* $\mathcal{E} = \{e_1, e_2, e_3, \dots\}$ *be the set of extractions. Our generative model will take* $\mathcal{E}$ *as input. An extraction item* e *is in the format of* $(f, p^{(*)}, o)$. *Here* o *stands for the observed frequency of fact tuples* f *that were extracted by pattern* $p^{(*)}$ *in* $\mathcal{E}$.

Definition 4 (Constraint). *Each commonsense rule (constraint) is represented as a variable. The variable is likely to be observed as 1. Examples:*
- *one* value *matches with only one* entity, *denoted as* $\mathcal{C}_{1v-1e}$ *that counts the number of such entities.*
- *one* entity *at one* time *matches with only one* value, *denoted as* $\mathcal{C}_{1(e,t)-1v}$ *that counts the number of values.*

2.2 Problem Definition

Suppose the set of extractions $\mathcal{E}$ have been obtained by pattern-based methods from text data. We define the problem as follows: **Given** a set of extractions $\mathcal{E}$, pattern sources $\mathcal{P}^{(*)}$, and the constraints $\mathcal{C}_a$ for attribute a, **infer** truth $\mathcal{T}$ for all temporal facts $\mathcal{F}$ contained in $\mathcal{E}$ and quality information for each pattern source $p^{(*)}$.

3 Proposed Approach

We mainly discussed the model detail of PGM with multiple Constraints $\mathcal{C}_{1(e,t)-1v}$ and $\mathcal{C}_{1v-1e}$, since it's the most complicated scenario while modeling constraint. The given input is the observed frequency of fact tuples extracted by a particular pattern and attached with a particular type of time signal. Figure 1 gives the plate notation of our model. Each node represent a variable. Blue nodes indicate hyper-parameter. Gray nodes stand for observable variable. And white nodes stand for latent variables we want to infer.

3.1 Generative Process

Our approach based on PGM is a generative process. We first generate a source. Next we generate a (entity, value, time)-tuple. Then we generate the
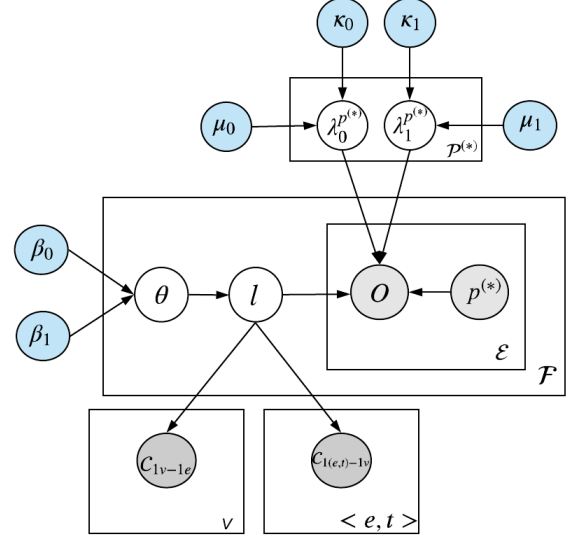

Figure 1: Probabilistic Graphic Model with Commonsense Constraint $\{\mathcal{C}_{1(e,t)-1v}, \mathcal{C}_{1(v)-1e}\}$

frequency based on the source reliability and the tuple's trustworthiness. Moreover, we generate variables according to the commonsense constraints. The variables counts the values/entities that can be matched to one entity/value with or without a time constraint (at one time) from the set of true tuples. The concrete meaning of each variable has been given in Table 1.

Temporal fact trustworthiness. For each temporal fact $f \in \mathcal{F}$, we first draw θ_f, i.e., the prior truth probability of fact f, from a *Beta* distribution with hyper-parameter β_0 and β_1:

$$\theta_f \sim Beta(\beta_0, \beta_1). \tag{1}$$

β_0 and β_1 represent the prior distribution of fact reliability. In practice, if we have a strong prior knowledge about how likely all or certain temporal facts are true, we can model it with the corresponding hyper-parameters. Otherwise, if we do not have a strong belief, we set a uniform prior, which means it's equally likely to be true or false, and our model can still infer the truth from other factors. After drawing the θ_f, we generate the truth label l_f from a *Bernoulli* distribution with parameter θ_f:

$$l_f \sim Bernoulli(\theta_f). \tag{2}$$

Pattern source reliability. As aforementioned, a reliable pattern source is more likely to extract true facts with higher counts, and extract false facts

Table 1: Symbols and their descriptions.

Symbol	Description
θ_f	[0, 1], trustworthiness of temporal fact tuple f
l_f	Boolean: label of temporal fact f
o_e	Integer: the observed frequency of fact f_e extracted by pattern $p_e^{(*)}$
$\lambda_0^{p^{(*)}}, \lambda_1^{p^{(*)}}$	Real numbers: reliability of pattern $p^{(*)}$ on giving false/true fact tuples
$\mathcal{C}_{1v-1e}$	Real number: the number of entities given one value v
$\mathcal{C}_{1(e,t)-1v}$	Real number: the sum of values given one entity e and one time t
Hyper-Parameter	
μ_0, μ_1	Integers: prior counts of false/true tuples extracted by a textual pattern
κ_0, κ_1	Integers: prior sums of false/true tuples extracted by a textual pattern
β_0, β_1	Integers: prior counts of false/true tuples

with lower counts. Therefore, we choose average count of false/true as latent pattern reliable weight, it's represented as $\lambda_0^{p^{(*)}}$, $\lambda_1^{p^{(*)}}$ for pattern $p^{(*)}$. The Gamma distribution is utilized because it is the conjugate prior of Poisson distributions. Initially, these two parameters are generated from *Gamma* distribution with hyper-parameter $\{\mu_0, \kappa_0\}/\{\mu_1, \kappa_1\}$, respectively. μ_0 and μ_1 represent the prior number of false/true fact the pattern extract, and κ_0 and κ_1 determine the prior sum of false/true fact count:

$$\lambda_0^{p^{(*)}} \sim Gamma(\mu_0, \kappa_0); \tag{3}$$

$$\lambda_1^{p^{(*)}} \sim Gamma(\mu_1, \kappa_1) \tag{4}$$

Extraction observation. For each extraction e $\in$ $\mathcal{E}$, it is composed of $\{f, p^{(*)}, o\}$. f_e denotes the temporal fact f belongs to e, $p^{(*)}$ denotes where it's extracted, o_e stands for extraction e's observation count. When the truth label of fact f_e is false, o_e is generated from *Poisson* distribution with $p^{(*)}$'s false speaking side parameter $\lambda_0^{p^{(*)}}$. While f_e is true, o_e is generated from Poisson Distribution with

$p^{(*)}$'s true speaking side parameter $\lambda_1^{p^{(*)}}$:

$$\begin{aligned} o_e &\sim Poisson(\lambda_0^{p^{(*)}}) \quad \text{if } l_{f_e} = 0, \\ o_e &\sim Poisson(\lambda_1^{p^{(*)}}) \quad \text{if } l_{f_e} = 1. \end{aligned} \tag{5}$$

Constraints. Finally, we draw the constraint variables. In temporal fact extraction, we define two variables $C_{1(e,t)-1v}$ and C_{1v-1e}. $C_{1(e,t)-1v}$ limits the number of truth on certain constraint key $\{e, t\}$. There are as many $C_{1(e,t)-1v}$ variables as unique $\{e, t\}$ keys:

$$C_{e,t} = \sum_f l_f, \quad f \in \mathcal{F}_{e,t}, \tag{6}$$

where $\mathcal{F}_{e,t}$ denotes a set of f with same $\{e, t\}$. Each $C_{1(e,t)-1v}$ is generated by $\mathcal{F}_{e,t}$ set.

C_{1v-1e} ilimits the truth of fact with same $\{v\}$:

$$C_v = \sum_{e \in E} l_{e,v} \quad \begin{cases} l_{e,v} = 1, & \text{if } \exists l_f = 1, \\ & f \in \mathcal{F}_{e,v}; \\ l_{e,v} = 0, & \text{otherwise.} \end{cases} \tag{7}$$

where $\mathcal{F}_v$ denotes set of fact with value v, $\mathcal{F}_{e,v}$ stands for a set of temporal fact f with same $\{e, v\}$, $\mathcal{F}_{e,v} \in \mathcal{F}_v$. $l_{e,v}$ denoted the truth label of v, e. Each $\mathcal{C}_{1v-1e}$ is generated by $\mathcal{F}_v$. If there is true fact $f \in \mathcal{F}_{e,v}$, then $l_{e,v}$ equals to one, otherwise, $l_{e,v}$ equal with zero.

4 Experiments

4.1 Dataset

We focus on attribute *country's president* and experiment on the same data set in the work of (Wang et al., 2019). It has 9,876,086 news articles (4 billion words) published from 1994–2010. We have 57,472 patterns, 116,631 temporal fact tuples, and 1,326,164 extractions. The dataset's ground truth was collected from Google and Wikipedia. It includes 3,175 true temporal facts of 130 countries.

4.2 Experiment Settings

4.2.1 Competitive methods

We compare our model with:
- TRUTHFINDER (Yin et al., 2008): It was a bootstrapping algorithm for structured data using C_{1v-1e}.
- LTM(Zhao and Han, 2012): It was a probabilistic model, assuming that the truth about an object contains more than one value. We set "object" as {entity, time} and set value as the temporal fact's value.

Table 2: Our proposed model performs better than baseline methods on finding temporal facts.

Method	Constraints e: Country; v: Person; t: year		Evaluation Setting			
	$\mathcal{C}_{1v-1e}$	$\mathcal{C}_{1(e,t)-1v}$	On (e,v,t)		On $(e,v,[t_{min},t_{max}])$	
			AUC	F1	AUC	F1
TRUTHFINDER		✗	0.0006	0.0012	0.0006	0.0012
LTM	✗	✗	0.1319	0.0199	0.2030	0.0218
LTM	✗	✔	0.0212	0.0505	0.0407	0.0793
TRUEPIE	✔	✗	0.0587	0.1430	0.0587	0.1430
MAJVOTE	✗	✔	0.3336	0.4318	0.4958	0.5927
TFWIN	✔	✔	0.4746	0.6361	0.5523	0.6489
Ours (PGMCC)	✗	✔	0.4840	0.6502	0.6006	0.7254
Ours (PGMCC)	✔	✔	**0.4987**	**0.6634**	**0.6075**	**0.7316**

• TRUEPIE (Li et al., 2018): It was a bootstrapping method using $\mathcal{C}_{1(v)-1e}$ and estimating pattern reliability.

• MAJVOTE (Goldman and Warmuth, 1995): It used the weighted majority voting strategy and returned the most frequent temporal fact.

• TFWIN (Wang et al., 2019): It was the state-of-the-art bootstrapping method for truth discovery on fact extraction. However, error propagation is serious in its iterative process.

4.2.2 Evaluation settings

All the methods can only find truth of temporal fact at one time point, e.g., (French, Jacques Chirac, 1995). However, due to the incompleteness of fact description in data, some time points of temporal facts could be missing. One way to improve the evaluation is to composite true temporal fact time points $\{e, v, t\}$ into temporal fact time period $\{e, v, [t_{min}, t_{max}]\}$. We evaluate the performance on both temporal fact time point $\{e, v, t\}$ and temporal fact time period $\{e, v, [t_{min}, t_{max}]\}$. To evaluate on time period $\{e, v, [t_{min}, t_{max}]\}$, we look at every single time points (e, v, t) in the period $(t \in [t_{min}, t_{max}])$.

4.2.3 Evaluation metrics

We evaluate all competitive methods using *precision*, *recall*, *F1 score*, and *AUC* (Area Under the Curve). Precision is the the fraction of temporal fact truth among all the temporal fact that were labelled as true. Recall is the fraction of true temporal facts our approach finds among the ground truth temporal facts. F1 score is the harmonic mean of precision and recall. For all of the metrics, higher score indicates that the method performs better.

4.3 Effectiveness

The results are given in Table 2. Our proposed method PGMCC consistently outperforms all the baselines on finding (country, president, time)-facts (i.e., presidential terms).

PGMCC vs LTM: PGMCC performs significantly better than LTM (+34.5% AUC; +64.4% F1) on evaluating time points, and performs better with (+40.45% AUC; +71.2% F1) on evaluating time periods. LTM was designed to solve structured truth finding like the bookseller example. So, there were many conflicts when applied to temporal fact extraction. PGMCC has multi-constraint as observable variables to alleviate the issue.

PGMCC vs TFWIN: PGMCC performs better than TFWIN (+2.4% AUC; +2.8% F1) on evaluating time points, and performs better with (+5.2% AUC; +8.3% F1) on evaluating time periods. TFWIN started with seed patterns and defined constraints as a rule to eliminate conflicting tuples. However, the inference on conflicts was based on local information (i.e., the current pattern reliability estimation). During this process, error might propagate through iterations. PGMCC is a probabilistic graphical model that can avoid error propagation by modeling constraints as variables and inferring truth with the global data distributions.

PGMCC with different constraints: See the last two rows in Table 2. For both PGMCC and TFWIN models, a complete constraint set, i.e., $\{\mathcal{C}_{1(v)-1e}$ and $\mathcal{C}_{1(e,t)-1v}\}$, gives the best performance. Partial constraint cannot fully identify conflicts or false tuples. $\mathcal{C}_{1(e,t)-1v}$ plays a significant role in extracting *country's president*.

Table 3: Pattern's reliability for country's presidency.

Textual Pattern p	$r_{p(post)}$	$r_{p(tag)}$
president `Person` of `Country`	0.920	0.870
`Country`'s current president `Person`,	0.978	0.250
`Country`'s newly elected president , `Person` ,	0.970	0.030
`Person`, now president of `Country`,	0.750	0.110
`Person`, who has ruled `Country`	0.438	0.994
$COUNTRY's former president `Person`	0.113	0.994
`Person`, who ruled `Country`	0.607	0.758
`Country` president `Person` signed	0.553	0.327
`Country` premier `Person`	0.012	0.010
`Country` foreign minister `Person`	0	0
`Country` golfer `Person`	0	0

4.4 Pattern Source Reliability Analysis

Table 3 presented some pattern examples and their scores. Here are our observations. First, the pattern "president `Person` of `Country`" is the only pattern that shows high reliability on both types of time signals (above 0.85). Second, the textual patterns that describe the current presidency are likely to have higher reliability on *text gen. time* than *temporal tag*, because the presidency was likely to be in the same time as the document was generated. These patterns usually have words such as "current", "newly", and "now". Third, the textual patterns that describe the past presidency are likely to have higher reliability on "tag" than "post", because the presidency was likely to be in the same time as the event (described in the sentence) happened but before the time of the document being generated. These patterns usually have words such as "have governed", "have ruled", "former", and "formerly".

5 Related Work

In this section, we review two relevant fields to our work, temporal fact extraction and truth discovery.

5.1 Truth Discovery

In big data era, the issue of "Veracity" on resolving conflicts among multi-source information is quite serious (Berti-Equille, 2015; Vydiswaran et al., 2011; Waguih and Berti-Equille, 2014; Dong et al., 2009; Galland et al., 2010; Xiao et al., 2016; Yin and Tan, 2011). Truth discovery methods find trustworthy information from conflicting multi-source (Xiao et al., 2015; Li et al., 2015). Several truth discovery methods have been proposed for various scenarios, and they have been successfully applied in diverse application domains. A few truth discovery methods are probabilistic model. LTM solved the "`Book`'s `author` list problem" and modeled its source in two-fold quality (Zhao et al., 2012). GTM solved the task of finding true numeric value of "New York City's population" (Zhao and Han, 2012). TEXTTRUTH found the true answer for a question from multi users (Zhang et al., 2018b).

5.2 Temporal Fact Extraction

Temporal fact extraction is to extract (entity, attribute name, attribute value)-tuples along with their time conditions from text corpora (Sil and Cucerzan, 2014; Hoang-Vu et al., 2016; Chekol, 2017; Zhang et al., 2018a; Shang et al., 2018; Zeng et al., 2019; Jiang et al., 2019). Textual patterns have been proposed to extract structured data from unstructured text data in an unsupervised way, such as E-A patterns (Gupta et al., 2014), parsing patterns (Nakashole et al., 2012), and meta patterns (Jiang et al., 2017). However, patterns are of different reliability and extractions are sometimes conflicting. In order to get reliable temporal fact, we addressed this problem using truth discovery.

6 Limitations and Future Work

Though the proposed approach show effectiveness in experiments, it and/or the study has several limitations. First of all, because collecting *temporal factual truth* for a variety of attributes is very expensive, in this study, we only studied a single relation type. In future work, we will apply the approach to other types of temporal-facts if correct constraints can be defined, such as sports team's players and spouse relationship. Second, though the patterns were generated by automated mining technologies

such as Meta Patterns (Jiang et al., 2017) (in other words, they are not hand-crafted), the pattern mining as a preprocessing step is needed. The approach is not end-to-end.

7 Conclusions

In this work, we proposed a probabilistic graphical model for inferring true facts and pattern reliability. It had two novel designs for temporal facts: (1) it modeled pattern reliability on temporal tag in text and text generation time; (2) it modeled commonsense constraints as observable variables. Experimental results demonstrated that our model outperformed existing methods.

Acknowledgements

References

Laure Berti-Equille. 2015. Data veracity estimation with ensembling truth discovery methods. In *Big Data (Big Data), 2015 IEEE International Conference on*, pages 2628–2636. IEEE.

Melisachew Wudage Chekol. 2017. Scaling probabilistic temporal query evaluation. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 697–706. ACM.

Luciano Del Corro, Abdalghani Abujabal, Rainer Gemulla, and Gerhard Weikum. 2015. Finet: Context-aware fine-grained named entity typing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 868–878.

Xin Luna Dong, Laure Berti-Equille, and Divesh Srivastava. 2009. Integrating conflicting data: the role of source dependence. *Proceedings of the VLDB Endowment*, 2(1):550–561.

Alban Galland, Serge Abiteboul, Amélie Marian, and Pierre Senellart. 2010. Corroborating information from disagreeing views. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 131–140. ACM.

Sally A Goldman and Manfred K Warmuth. 1995. Learning binary relations using weighted majority voting. *Machine Learning*, 20(3):245–271.

Rahul Gupta, Alon Halevy, Xuezhi Wang, Steven Euijong Whang, and Fei Wu. 2014. Biperpedia: An ontology for search applications. *Proceedings of the VLDB Endowment*, 7(7):505–516.

Tuan-Anh Hoang-Vu, Huy T Vo, and Juliana Freire. 2016. A unified index for spatio-temporal keyword queries. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 135–144. ACM.

Meng Jiang, Jingbo Shang, Taylor Cassidy, Xiang Ren, Lance M Kaplan, Timothy P Hanratty, and Jiawei Han. 2017. Metapad: Meta pattern discovery from massive text corpora. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 877–886. ACM.

Tianwen Jiang, Tong Zhao, Bing Qin, Ting Liu, Nitesh V Chawla, and Meng Jiang. 2019. The role of" condition" a novel scientific knowledge graph representation and construction model. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1634–1642.

Qi Li, Meng Jiang, Xikun Zhang, Meng Qu, Timothy P Hanratty, Jing Gao, and Jiawei Han. 2018. Truepie: Discovering reliable patterns in pattern-based information extraction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1675–1684. ACM.

Yaliang Li, Qi Li, Jing Gao, Lu Su, Bo Zhao, Wei Fan, and Jiawei Han. 2015. On the discovery of evolving truth. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 675–684. ACM.

Ndapandula Nakashole, Gerhard Weikum, and Fabian Suchanek. 2012. Patty: a taxonomy of relational patterns with semantic types. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 1135–1145. Association for Computational Linguistics.

Nils Reimers, Nazanin Dehghani, and Iryna Gurevych. 2016. Temporal anchoring of events for the timebank corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 2195–2204.

Jingbo Shang, Jialu Liu, Meng Jiang, Xiang Ren, Clare R Voss, and Jiawei Han. 2018. Automated phrase mining from massive text corpora. *IEEE Transactions on Knowledge and Data Engineering*, 30(10):1825–1837.

Avirup Sil and Silviu-Petru Cucerzan. 2014. Towards temporal scoping of relational facts based on wikipedia data. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning*, pages 109–118.

VG Vydiswaran, ChengXiang Zhai, and Dan Roth. 2011. Content-driven trust propagation framework. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 974–982. ACM.

Dalia Attia Waguih and Laure Berti-Equille. 2014. Truth discovery algorithms: An experimental evaluation. *arXiv:1409.6428*.

Xueying Wang, Haiqiao Zhang, Qi Li, Yiyu Shi, and Meng Jiang. 2019. A novel unsupervised approach for precise temporal slot filling from incomplete and noisy temporal contexts. In *The World Wide Web Conference*, pages 3328–3334. ACM.

Houping Xiao, Jing Gao, Qi Li, Fenglong Ma, Lu Su, Yunlong Feng, and Aidong Zhang. 2016. Towards confidence in the truth: A bootstrapping based truth discovery approach. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1935–1944. ACM.

Houping Xiao, Yaliang Li, Jing Gao, Fei Wang, Liang Ge, Wei Fan, Long H Vu, and Deepak S Turaga. 2015. Believe it today or tomorrow? detecting untrustworthy information from dynamic multi-source data. In *Proceedings of the 2015 SIAM International Conference on Data Mining*, pages 397–405. SIAM.

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

Xiaoxin Yin and Wenzhao Tan. 2011. Semi-supervised truth discovery. In *Proceedings of the 20th international conference on World wide web*, pages 217–226. ACM.

Qingkai Zeng, Mengxia Yu, Wenhao Yu, Jinjun Xiong, Yiyu Shi, and Meng Jiang. 2019. Faceted hierarchy: A new graph type to organize scientific concepts and a construction method. In *Proceedings of the Thirteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-13)*, pages 140–150.

Chao Zhang, Fangbo Tao, Xiusi Chen, Jiaming Shen, Meng Jiang, Brian Sadler, Michelle Vanni, and Jiawei Han. 2018a. Taxogen: Constructing topical concept taxonomy by adaptive term embedding and clustering. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.

Hengtong Zhang, Yaliang Li, Fenglong Ma, Jing Gao, and Lu Su. 2018b. Texttruth: an unsupervised approach to discover trustworthy information from multi-sourced text data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2729–2737. ACM.

Bo Zhao and Jiawei Han. 2012. A probabilistic model for estimating real-valued truth from conflicting sources. *Proc. of QDB*.

Bo Zhao, Benjamin IP Rubinstein, Jim Gemmell, and Jiawei Han. 2012. A bayesian approach to discovering truth from conflicting sources for data integration. *Proceedings of the VLDB Endowment*, 5(6):550–561.