Learning a Typhoon Bayesian Network Structure from Natural Language Reports

Zhangrui Yao, Junhan Chen, Yinghui Pan, Yifeng Zeng, Biyang Ma, Zhong Ming

Learning a Typhoon Bayesian Network Structure from Natural Language Reports*

Zhangrui Yao[1], Junhan Chen[1], Yinghui Pan*[2], Yifeng Zeng*[3], Biyang Ma[3], and Zhong Ming[2]

[1] Department of Automation and Xiamen Key Laboratory of Big Data Intelligent Analysis and Decision-making, Xiamen University, China
[2] College of Computer Science and Software Engineering, Shenzhen University, China
[3] Department of Computer and Information Sciences, Northumbria University, UK

Abstract. Given the huge toll caused by natural disasters, it is critically important to develop an effective disaster management and emergency response technique. In this article, we investigate relationships between typhoon-related variables and emergency response from natural language (NL) reports. A major challenge is to exploit typhoon state information for typhoon contingency plan generation, especially from unstructured text data based on NL input. To tackle this issue, we propose a novel framework for learning typhoon Bayesian network structures (FLTB), which can extract typhoon state information from unstructured NL, mine inter-information causal relationships and then generate Bayesian networks. We first extract information about typhoon states through NL processing (NLP) techniques, and then analyze typhoon reports by designing heuristic rules to identify causal relationships between states. We leverage these features to improve the learned structures and provide user-interaction mechanisms to finalize Bayesian networks. We evaluate the performance of our framework on real-world typhoon datasets and develop the Bayesian networks based typhoon emergency response systems.

Keywords: Causal Structure Learning · Bayesian Networks · Typhoon Emergency Plan

1 Introduction

Intelligent decision models have shown significant application values in the prediction of natural disaster events [4]. An effective decision-making model can be built by analyzing how natural disasters occur and what they are leading to. The analysis is often accompanied with a large amount of text, such as news reports and real-time microblogs.

In many cases, emergency response systems often need to interact with users lacking professional knowledge. Hence, probabilistic graphical models (PGMs)

* Supported by NSF: 62176225 and 61836005. Yinghui and Yifeng are the corresponding authors with the emails: panyinghui@szu.edu.cn and yifeng.zeng@northumbria.ac.uk.

are an ideal tool for developing decision models in emergency response systems [8]. In this paper, we construct Bayesian network as the typhoon decision model. Due to its high interpretability, Bayesian networks have become a reliable decision-making tool in many application fields.

There have been many studies on Bayesian network structure learning [2], but existing techniques rarely focus on unstructured data. In addition, the information in the text is highly fragmented, the text features are sparse, and the available open annotated corpus is scarce. Consequently, it is rather difficult to directly generate a Bayesian network from the text.

Under this setting, we propose a Bayesian network structure learning scheme based on natural language inputs, namely FLTP. We adopt Bert model to extract the state pairs of typhoons in which its bidirectional encoder can extract feature factors of each word. Then we employ domain knowledge to normalize the state of the variables. We design a rule template to find causal relationships between states and build the Bayesian network. We also provide a knowledge-based approach to pruning Bayesian networks. The resulting Bayesian networks can be compiled by a general PGM tool, e.g. GeNIe[1] or HUGIN[2] application. In addition, we design an interactive component so that users can provide useful knowledge and optimize the final output of the Bayesian networks.

The rest of the article is organised as follows. Section 2 reviews related works. Section 3 presents our novel techniques that learn Bayesian network structures in a typhoon contingency plan. Experiments are conducted in Section 4. Section 5 concludes our work and discusses directions of future research.

2 Related Works

An intelligent decision model has many application precedents in the field of disaster response [8]. Most of the existing typhoon emergency decision-making research is based on statistical methods, and the data affecting emergency decision-making is obtained through demographic characteristics such as questionnaires [11]. However, these statistics-based models also have many limitations, such as slow data collection process and analysis process. Although learning Bayesian networks from data has been extensively studied in various fields [1],learning network structure from text input remains a challenge. Learning Bayesian network structure can be divided into two steps - states and their relations extraction.

There are few studies on extracting node status directly from text, and most of them are extracted by rules. For example, Trovati et al. [10] extracted and populate BN fragments based on grammar and lexical properties. In recent years, many studies have applied natural language processing technology in the fields of disaster prevention [3]. However, most of the current works only identify and analyze disaster events without extracting their relations.

The most noteworthy information in the text is the causal relationship, which contains a lot of knowledge and resources in the form of causal relationship. Understanding possible causal relationships between pairs of events is of great

[1] https://www.bayesfusion.com/genie/
[2] https://www.hugin.com/

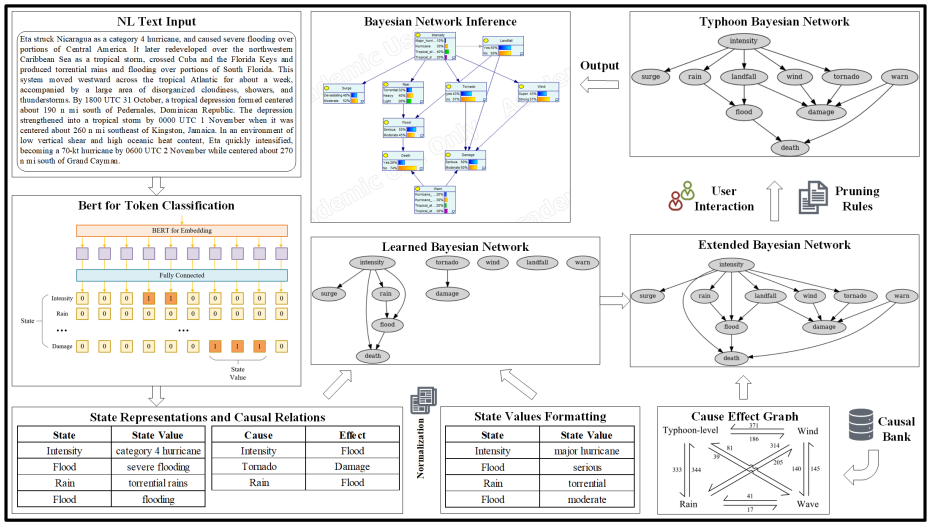

Fig. 1. A framework of learning typhoon Bayesian network structures (FLTB) contains five main parts from learning typhoon information from NL text to construct a Bayesian network.

significance in Event Prediction and Decision Processing [7]. The methods of causal extraction can be divided into rule-based methods [9], machine learning-based methods[12], and deep learning-based methods [5]. In this article, we adapt the machine learning based methods with additional rule-based knowledge model to discover causal relations in Bayesian networks.

3 The Framework of Learning Typhoon Bayesian Network Structures

FLTB focuses on the identification of states and relations, and then learn the network structure. As shown in Fig. 1, FLTB contains five parts. ① State extraction. ② State values Normalization. ③ Causal relationship extraction. ④ Bayesian network Generation ⑤ Network output and inference.

3.1 State Extraction Model

The goal of state extraction is to identify all possible (state, state value) pairs in a sentence. We adapt BERT(Bidirectional Encoder Representations from Transformers) for token classification in the FLTB. In general, for a given data set D with utterance x_j and state pair label $T_j = (s, v)$, we aim to maximize the probabilistic likelihood in the data set D: $p_\theta(s, v|x) = \prod_{j=1}^{|D|}[\prod_{(s,v)\in T_j} p((s, v)|x_j)] = \prod_{j=1}^{|D|}[\prod_{s\in T_j} p(v|x_j)]$.

The state value tagger optimizes the following likelihood function to identify the span of value v given a sentence representation x: $p_{\theta s}(v|x) = \prod_{i=1}^{L}(p_i^s)^{I\{y_i^s=1\}}(1-p_i^s)^{I\{y_i^s=0\}}$,where L is the length of the sentence. $I\{z\} = 1$ if z is true and 0 otherwise. y_i^s is the binary tag of state for the i^{th} token in x.

3.2 Standardize State Information

Due to the diversity and complexity of NL expressions and the characteristics of the extraction model, the state we capture may have different representations. Thus, we need to merge and unify the extracted state information.

We create a normalization base to standardize the states we extract. Specifically, we screen out representative statements from typhoon news on the Internet, and summarize their representative feature representations. We develop a uniform representation of state values. For example, we filter out some representative expressions of *rainfall* from typhoon news: *widespread heavy rain, local heavy rain*, *general heavy rain*, etc. Then we unify all these expressions into the state value of *heavy rain* in the state *rain*. Second, for the numerical status value attributes obtained in the news, we design rules to classify them into levels. After this process, we get the normalized state values.

3.3 Causal Relationship Extraction

We Identify causality by checking whether a sentence contains causal patterns. Then we use a Bayesian classifier to filter out noisy pairs.

3.3.1 Identify Causal Patterns

To represent the syntactic structure of the causal relationship in the sentence, we define some lexical syntactic patterns based [9].

- Using causal links to link two phrases or sentences. (e.g. therefore, because)
- Using a single verb that expresses cause and effect. (e.g. cause, produce)
- Using causal verbs in the passive voice. (e.g. cause by, triggered by)
- Using a single preposition to connect cause and effect (e.g. after, from)

To make the causal patterns more general, we expand the list by adding common phrases that contain those words. For example, the *cause* in the cause and effect list can be extended to include phrases such as *the main causes of*, *the leading cause of*.

For each causal pattern, a corresponding regular expression is defined to identify sentences containing this pattern, e.g. the regular expression $(.*) < produce|cause|triggers|... > (.*)$ can match the type of causative verbs. Then we define a series of rules to confirm cause and effect.

3.3.2 Causality Extraction Rules

The rules for detecting causality are based on syntactic structures. We divide the causal words into forward causal words (e.g. cause, therefore) and reverse causal words (e.g. resulted by, caused by), and then determine which part is the cause and which the effect based on the position of the state and the trigger words in the text. The main rule is:

$$((C, P^+) \cup (P^-, C)) \cap ((E, P^-) \cup (P^+, E)) \rightarrow cRel(C, E) \quad (1)$$

where P^+ and P^- are forward causal and reverse causal respectively, and (C, P^+) means that the state C is in the front of P^+.

In order to deal with multiple causal relationships in the same sentence, we also study causal chain relationships in a sentence. For example, a sentence contains a sequence of causal events, event e_1 causes another event e_2 to occur, which in turn causes event e_3 to occur. In this case, e_2 would be marked as an *effect* for $e1$ and a *cause* for e_3.

Table 1. Main notations used in this paper.

x	State
X_T	The state set of typhoon information
X_M	The state set of meteorology information
X_S	The state set of secondary disaster
X_D	The state set of disaster information
(a, b)	a and b form a causal relationship, with a as the cause and b as the effect
$X_h > X_l$	X_h has a higher priority than X_l

Using the above rules, we create a causal rule base. We use a dependency parser to segment each input sentence into words. Once any causal marker words are identified in the sentence, we apply the above rules to them, thereby extracting causal relationships between states.

3.3.3 Candidate Pair Filtering

The relationships we extracted using the above rules may be noisy, there are some pairs (C, E) that do not have a causal relationship. To address this, we create a causal database of 8392 sentences from the collected sentences (among them, 1541 are marked as causality), and then obtain prior knowledge from the causal database and use a Naive Bayes classifier to remove false causal pairs.

$$P(c_i|r) = P(r|c_i) * P(c_i)/P(r) \tag{2}$$

where, with $i = 0$ or 1, c_1 is causal and c_0 is non-causal, and r is candidate causal pair, $P(c_1)$=1541/8392, $P(c_0)$=6851/8392. $P(r|c_i)$ refers to the number of r in c_i. According to the Bayesian classification rule, the relation is classified as causal if $P(c_1|r) > P(c_0|r)$.

3.4 Generate Typhoon Bayesian Network

3.4.1 Generate Backbone Network

Firstly, we consider aspects of the impact caused by typhoons, and classify states into four categories based on catastrophology [13]: ① Typhoon information: The properties of the typhoon itself, like *intensity*, *landfall*... ②Meteorological information: Typhoon would bring some meteorological impacts, like *rain*, *wind*, *surge*... ③ Secondary disasters: Typhoons would cause a series of secondary hazards, like *tornado*, *flood*... ④ Disaster information: The impact of the typhoon and the measures taken by the government, including *damage*, *death*, *warn*... We define the main notations in Table 1.

3.4.2 Prior Knowledge-based Bayesian Network Extension

We construct a prior causal knowledge base through the causal bank [6]. Then the Bayesian network is extended through this prior knowledge base and typhoon domain rules we develop below.

Rule 1: We collect the information of typhoon disaster chain and combine it with the knowledge of disaster science [14]. We assume the following priority of the four categories: $X_T > X_M > X_S > X_D$. The lower ranked states can not be used as the cause of the higher states.

Rule 2: We find that issuing the wrong level of warning would lead to more severe disasters. So we assume that the *warn* status only can be the reason for the others in disaster information status.

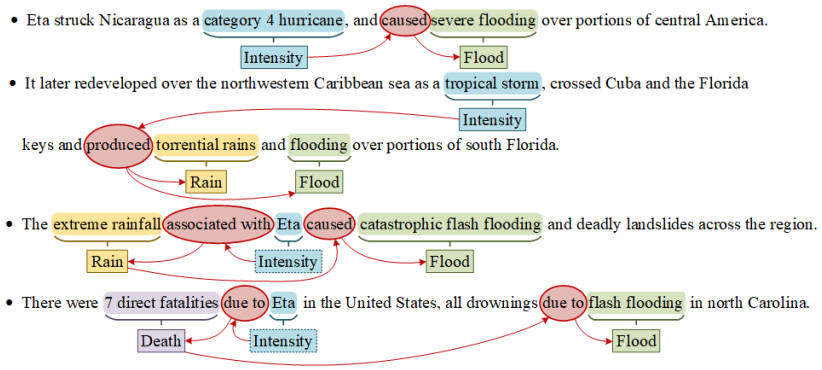

Fig. 2. Results of extracting the S1-S4 sentences.

Rule 3: Since the subject of our study is typhoon, the secondary disaster should be caused by typhoon. Therefore we assume that states of the secondary disaster must have at least one cause node.

Thus, for isolated nodes in the backbone network, we extend the network through selecting the node with the highest frequency in the prior knowledge base to form causal relationships.

3.4.3 Bayesian Network Pruning

Based on the nature of the network structure and the pattern of typhoon event development, we develop the following rules for pruning a network structure.

Rule 1: We assume that there is causal transitivity between the four typhoon categories. In other words, The states between four levels follow the head-to-tail, i.e., $A \rightarrow B \rightarrow C$. A and C are independent given B.

Rule 2: We find when there are multiple weather states pointing to the same secondary disaster information, the relationship between them is irrelevant and can be pruned.

Rule 3: We assume that the secondary hazards are all caused by typhoons. No causal relationship will be formed between the secondary hazards, i.e., secondary disaster states can not form a causal relationship with each other.

In addition, we introduce user interaction in practical applications so that users can define specific causal relationships according to the actual situation.

3.4.4 Bayesian Network Output and Inference

In the last step, we create a Bayesian network using states and causal relationships obtained by FLTB. Then we use PGM software (e.g. GeNIe) to inference the network. Having a Bayesian network can enhance many downstream applications, including question answering and inference.

4 Experimental Results

We collected 169 influential hurricane reports from the U.S. National Hurricane Center that occurred between 2011 and 2021 . We create the training data from given sentences for extracting states, and then test FLTB on learning Bayesian

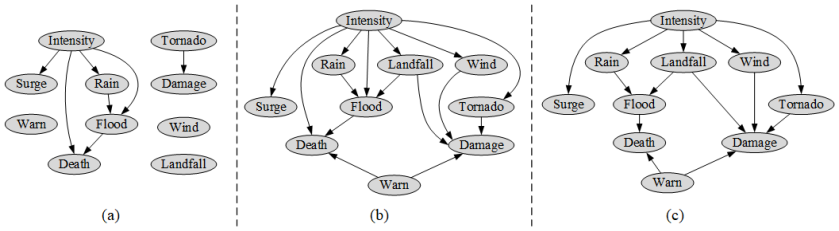

Fig. 3. Results of the step of generating a typhoon Bayesian network through FLTB. (a) The basic Bayesian network, (b) The extended Bayesian network, and (c) The final output Bayesian network.

network from news of the typhoon Eta, in order to illustrate that our FLTB is effective and automatic in unstructured data.

We have built the FLTB based typhoon BN learning engine and developed the demo as shown in the link [1].

Table 2. The states obtained from sentences S1-4 by our FLTB.

Sentence	States and State Values	States Normalization	Causal Relationships
S1	(Intensity , category 4 hurricane) (Flood , severe flooding)	(intensity , major hurricane) (flood , serious)	(intensity , flood)
S2	(intensity , tropical storm) (rain , torrential rains) (flood , flooding)	(intensity , tropical storm) (rain , torrential) (flood , moderate)	(intensity , rain) (intensity , flood)
S3	(rain , extreme rainfall) (flood , catastrophic flash flooding) (intensity , eta)	(rain , torrential) (flood , serious) (intensity , eta)	(intensity , rain) (rain , flood)
S4	(flood , flash flooding) (death , 7 direct fatalities) (intensity , eta)	(flood , serious) (death , serious) (intensity , eta)	(intensity , death) (flood , death)

Taking the four input sentences as an example, Fig. 2 illustrates the extracting results of sentences S1, S2, S3 and S4, respectively.

After an automatic process on the dataset, the sets of states, state values and causal relationships, corresponding nodes, random variables and directed edges are shown in Table 2. As shown in Fig. 3 (a), FLTB can construct a basic Bayesian network based on the extracted information. Then we clearly see that the backbone network is incomplete with lots of isolated nodes . We add rule constraints to extend isolated nodes based on the external database and the causal bank.

From the new network in Fig. 3 (b), we can see that the network is obviously complex and redundant. Hence we simplify the network according to the constrains formulated in Section 3.5. The result in Fig. 3 (c) illustrates the automation and rationality of FLTB.

5 Discussions and Conclusions

We develop a novel framework, namely FLTB, to automate the structure learning in a Bayesian network from the inputs of NL text. This is also the first attempt at

[1] https://github.com/lamingic/LBfT

extracting typhoon states for constructing Bayesian network by using the NLP technique. The new method can see the benefit of generating Bayesian networks for small unstructured text data as shown in our empirical study.

As we are deploying this framework in a practical application of generating typhoon contingency plans, there is a need to further reduce the user interaction and develop a more automatic process. In addition, we may consider adopting disaster news reports from multiple disasters as input to gradually establish a larger network, and build more accurate plans.

References

1. Behjati, S.: An order-based algorithm for learning structure of bayesian networks. In: International Conference on Probabilistic Graphical Models, 11-14 September 2018 (2018)
2. Constantinou, A.C.: Learning bayesian networks with the saiyan algorithm. ACM Transactions on Knowledge Discovery from Data (TKDD) **14**(4), 1–21 (2020)
3. Domala, J., Dogra, M., Masrani, V., Fernandes, D., D'souza, K., Fernandes, D., Carvalho, T.: Automated identification of disaster news for crisis management using machine learning and natural language processing. In: 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). pp. 503–508. IEEE (2020)
4. Huo, Y., Tang, J., Pan, Y., Zeng, Y., Cao, L.: Learning a planning domain model from natural language process manuals. IEEE Access **8**, 143219–143232 (2020)
5. Li, Z., Li, Q., Zou, X., Ren, J.: Causality extraction based on self-attentive bilstm-crf with transferred embeddings. Neurocomputing **423**, 207–219 (2021)
6. Li, Z., Ding, X., Liu, T., Hu, J.E., Van Durme, B.: Guided generation of cause and effect. arXiv preprint arXiv:2107.09846 (2021)
7. Miranda Ackerman, E.J.: Extracting a causal network of news topics. In: OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". pp. 33–42. Springer (2012)
8. Sankar, A.R., Doshi, P., Goodie, A.: Evacuate or not? a pomdp model of the decision making of individuals in hurricane evacuation zones. In: Uncertainty in Artificial Intelligence. pp. 669–678. PMLR (2020)
9. Sorgente, A., Vettigli, G., Mele, F.: Automatic extraction of cause-effect relations in natural language text. DART@ AI* IA **2013**, 37–48 (2013)
10. Trovati, M., Hayes, J., Palmieri, F., Bessis, N.: Automated extraction of fragments of bayesian networks from textual sources. Applied Soft Computing **60** (2017)
11. Yongsatianchot, N., Marsella, S.: Modeling human decision-making during hurricanes: From model to data collection to prediction. In: AAMAS Conference proceedings (2019)
12. Zhao, S., Liu, T., Zhao, S., Chen, Y., Nie, J.Y.: Event causality extraction based on connectives analysis. Neurocomputing **173**, 1943–1950 (2016)
13. Zheng, L., Wang, F., Zheng, X., Liu, B.: A distinct approach for discovering the relationship of disasters using big scholar datasets. In: International Conference on Geo-Spatial Knowledge and Intelligence. pp. 271–279. Springer (2017)
14. Zheng, L., Wang, F., Zheng, X., Liu, B.: Discovering the relationship of disasters from big scholar and social media news datasets. International Journal of Digital Earth **12**(11), 1341–1363 (2019)