Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(4): 7217–7233.
DOI: 10.3934/mbe.2023312
Received: 16 November 2022
Revised: 09 January 2023
Accepted: 22 January 2023
Published: 10 February 2023

Research article

Temporal fact extraction of fruit cultivation technologies based on deep learning

Xinliang Liu[1,2,3], **Lei Ma**[2,3,*], **Tingyu Mao**[2,3] **and Yanzhao Ren**[4]

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
[2] School of E-business and Logistics, Beijing Technology and Business University, Beijing 100048, China
[3] National Engineering Laboratory for Agri-product Quality Traceability, Beijing Technology and Business University, Beijing 100048, China
[4] School of Computer Science and Engineering, Beijing Technology and Business University, Beijing 100048, China

* **Correspondence:** Email: malei@st.btbu.edu.cn.

Abstract: There are great differences in fruit planting techniques due to different regional environments. Farmers can't use the same standard in growing fruit. Most of the information about fruit planting comes from the Internet, which is characterized by complexity and heterogeneous multi-source. How to deal with such information to form the convenient facts becomes an urgent problem. Information extraction could automatically extract fruit cultivation facts from unstructured text. Temporal information is especially crucial for fruit cultivation. Extracting temporal facts from the corpus of cultivation technologies for fruit is also vital to several downstream applications in fruit cultivation. However, the framework of ordinary triplets focuses on handling static facts and ignores the temporal information. Therefore, we propose Basic Fact Extraction and Multi-layer CRFs (BFE-MCRFs), an end-to-end neural network model for the joint extraction of temporal facts. BFE-MCRFs describes temporal knowledge using an improved schema that adds the time dimension. Firstly, the basic facts are extracted from the primary model. Then, multiple temporal relations are added between basic facts and time expressions. Finally, the multi-layer Conditional Random Field are used to detect the objects corresponding to the basic facts under the predefined temporal relationships. Experiments conducted on public and self-constructed datasets show that BFE-MCRFs achieves the best current performance and outperforms the baseline models by a significant margin.

Keywords: temporal facts; information extraction; fruit cultivation technologies; deep learning; temporal expression

1. Introduction

Fruit cultivation is an important part of the food industry. There are a number of uniform standards in planting techniques. However, due to the great differences in environmental conditions in different regions, it is impossible to use a unified standard for fruit cultivation. Moreover, the information about fruit cultivation on different platforms are complicated and not reusable. The information extraction is the most essential technique for extracting knowledge, and it performs well within the task of handling unstructured and complex data [1]. An outsized range of large-scale knowledge bases, like Database Pedia (DBpedia), Never-Ending Language Learner (NELL) [2], Probase (Probabilistic Taxonomy) [3] and YAGO (A Multilingual Knowledge Base) [4] have been constructed by entity-relationship extraction technique, which contain millions of entities and facts. The majority of studies on methods of information extraction have focused on the acquisition of basic (static) facts while not considering the constraints of temporal elements. However, the sequential relationships implied by temporal information are of great importance [5], especially in the field of agronomy. In this field, temporal information is an elementary component of information like Date of pollination, Days of growth, Date of fertilization, etc. A large number of facts are only valid in the specified time interval. "Barack Obama President of United States" is a correct fact, but not valid anymore. "Apples can be stored at room temperature" is a simple fact, however it's solely helpful knowledge if it's a storage time hooked up to that. Therefore, extracting temporal information from unstructured text is a vital task in the field of fruit cultivation.

Researches on temporal information extraction could be divided into three main directions according to task boundaries: temporal expression extraction, temporal relation extraction and event-time extraction [6]. The mode of temporal expression extraction aims at finding all temporal descriptions in the text and normalizing them in the same way. The temporal relation extraction focuses on ordering the trigger words of the events and extracting the temporal relations between event pairs. In event-time extraction, the time of the event is the extraction target. Table 1 shows the example of three extraction directions.

Table 1. The examples of three extraction tasks. The temporal expression is marked in orange and the event is marked in blue.

Boundary	Sentence	Target
Temporal Expression	I need three hours to finish the homework.	three hours
Temporal Relation	Apples can be stored for 3 days after ripening.	After (ripen, store)
Event-time Expression	JinMei kiwifruit passed the national variety validation in 2014.	validation, 2014

In the first extraction task, Cao et al. [7] proposed XLTime, a novel framework for multilingual temporal expression extraction, which transferred English to other language to extract the target temporal expression. Many scholars focus on event-time expression extraction, Ling and Weld [8] first recognized the event and time expressions in each sentence with Evita (the event recognizer) and GUTime (time expression recognizer), and then applied dependency trees and role annotators to extract event and time attributes and dependency features. Previous studies abstracted the event into a verb or noun to extract its corresponding temporal expression. This could not extract the detailed temporal facts in a complete way. According to the corpus characteristics in this paper, this study simplifies and defines the temporal information extraction task as temporal fact extraction. The framework of temporal fact extraction is represented as ((s, r_b, o), r_t, t). s and o are the head and tail entities of the basic facts, respectively. r_b is the relationship between s, o. t represents the time expression and r_t is the connection of the basic fact to t. The basic facts are extracted first, and then the temporal expressions of the basic facts are extracted. In this paper, we make the following contributions:

- A general framework Basic Fact Extraction and Multi-layer CRFs (BFE-MCRFs) is constructed to extract temporal information from simple unstructured text, and it could use multi-layer Conditional Random Field (CRFs) to deal with single entity overlap.
- The links between basic facts and time expressions are defined as relations, including Start Time (ST), End Time (ET) and Lasting Time (LT).
- A dataset of temporal knowledge of fruit cultivation technologies (FCTD) in Chinese is established, which could be used by subsequent scholars studying temporal extraction in Chinese.

2. Related Works

Temporal information extraction is a vital task in knowledge extraction [9], which aims to identify the temporal representation of facts from a bunch of complicated texts. Temporal information can be used to order events and measure the durations of events. Li et al. [10] presented a temporal information extraction system for Chinese temporal events from financial news. The ranking of events emerged as an essential task, which was helpful to find all feasible times when a given event may occur. The temporal IE system is applied to deal with the task of question-answering in event-based queries, temporal queries and complex queries. Tanev et al. [11] presented a real-time news event extraction system developed by the Joint Research Centre of the European Commission. Strötgen et al. [12] designed a system combining temporal and geographical information for ordering events and spatial information. All of these researchers developed event ordering systems by extracting temporal information, and they manually defined extraction constraints and manners based on linguistic knowledge.

Three mainstream temporal information extraction task methods are analyzed below, which are the rule-based method, machine learning method and hybrid method. These methods have an inspiring effect on the design of BFE-MCRFs model.

2.1. Rule-based methods

Research on temporal information extraction has become a popular topic [13], and scholars initially considered using artificially defined rules or frameworks for extraction. Mani et al. [14] extracted pre-defined feature words such as tense, temporal prepositions, and discourse connectives from the corpus TimeBank [15], which is a corpus of 183 English news documents that have been manually annotated according to TimeML (the annotation specification of the ISO Time Expression Annotation Language). By using rule-based extraction methods, researchers have constructed several knowledge bases with time-stamped facts, like T-Yago [16], WikiData [17] and Yago2 [18]. These knowledge bases try generic expressions to extract temporal facts from semi-structured data, but they have limitations in the application of unstructured text.

Some studies have suggested pattern-based extraction methods in order to extract temporal expressions from semi-structured and unstructured texts. First, a temporal pattern is generated using rules or self-defined constraints, and then the corpus is processed using the temporal pattern to obtain temporal facts. Kuzey and Weikum [19] focused on harvesting temporal facts and events from Wikipedia to create a temporal ontology. They defined the manually crafted rules and the types of relationships to extract temporal information from semi-structured data such as infoboxes, lists, and the titles of Wikipedia articles. Then a quaternary pattern that already existed in the knowledge base was applied to free texts to create new temporal facts. This extraction method was not flexible enough due to its high relevance to the corpus. Liu et al. [20] designed the temporal pattern with the classical techniques, including corpus annotation, pattern generation, scoring and clustering. This method sought the most relevant indicating phrase for each predefined relationship. Then the collection of temporal patterns was formed based on the calculated candidate patterns, which were finally clustered and analyzed to obtain the final temporal patterns. The top-k temporal patterns were designed to extract the temporal instances.

Rule-based and pattern-based approaches are extremely effective in extracting facts from a specific structure. However, the design of rules and patterns is very laborious and time-consuming, even if there is only one relationship. And these approaches have very little generalization ability. Moreover, the representation of unstructured text is usually vague, which makes the extraction of temporal patterns very difficult.

2.2. Machine learning and hybrid methods

Based on machine learning, hybrid methods are attempted to be applied to temporal information extraction. A hybrid model incorporating rules-based on CRF and Support Vector Machines (SVM) was investigated by Tang et al. [21] to extract temporal events, temporal expressions, and event features. This model performed pretty well in the clinical domain. Moharasan and Ho [22] presented a novel semi-supervised framework, which exploited the Naive Bayes Classifier. They identified the potential generated candidate pairs that helped to increase the performance of the relation classification between events or expressions in clinical narratives. Machine learning-based approaches perform effectively on small-scale text corpora, but the effectiveness of the model depends heavily on the selection of complex features and the quality of the annotated corpus. Deep learning-based methods can be well suited to the task of automatic feature extraction. Han et al. [23] first built a neural structured prediction model with joint representation learning to make predictions on events-temporal and relations simultaneously. Then he designed a framework [24] that enhanced deep neural networks with distributional constraints constructed by probabilistic domain knowledge. These authors pay attention to improving the efficacy of temporal relation extraction.

The above studies focus on the extraction of either temporal relations [25–27] or temporal expressions [28, 29] and these methods perform significantly well. However, the temporal information describes the details of facts. Separating these two tasks could not extract the temporal facts that construct the temporal knowledge graph.

Based on above researches, this paper proposes a joint model named BFE-MCRFs to extract temporal facts in the field of fruit cultivation. BFE-MCRFs attempts to extract temporal expressions by feeding basic facts into a high-level model, which is a general framework for extracting temporal facts from large-scale unstructured text. Moreover, BFE-MCRFs performs very well in the task of single entity overlapping extraction that makes it crucially different from previous works.

3. Materials and Methods

3.1. Temporal Fact Defination

This paper defines the extraction task as higher order facts extraction issue, which tries the schema ((s, r_b, o), r_t, t) to extract time expression.

3.1.1. Basic fact

First order facts are modeled as basic facts, which describe the major elements of the sentence, including subject, relation and object. An example fact in the sentence “Kiwifruit is rich in folic acid”, (kiwifruit, enriched, folic acid) is an instance of the binary relation and the subject and the object are entities.

3.1.2. Temporal expression boundary

Temporal expression varies from field to field. Texts like newswire text or clinical narratives have precise point-in-time descriptions, e.g., a news event occurs on May 1, 2022, at 5:00 pm. The Wikipedia text describes the temporal information as a specific date, like “2022-5-1” and “Christmas 2022”. However, the time expressions in the corpus of this paper are expressed in a timespan or vague time, such as “late April, early May” and “around 15 days.” Hence, this study aims to extract the time expression in the text. Determining whether it is a precise or ambiguous temporal expression is out of the scope of this work. As a result of the analysis of the text corpus, it is concluded that the following three temporal categories are included:
Temporal duration the time points of the sentence. It clearly defines the point in time, for instance, 30 days, a week.
Temporal interval the timespan of the sentence. It includes the start time and end time of the occurrence of the basic facts, for example (late April, early May).
Temporal prepositional phrase the prepositional phrase expressions of the sentence. It expresses the chronological order of occurrence of two entities. The temporal prepositional expressions are instantiated as prepositional relation types in the content of this study. Such as, in the sentence “Yate can be stored for 3 days after ripening.”, which contains the temporal fact (store, after, ripen, 3 days).

3.1.3. Temporal fact

A higher order fact usually contains the basic facts and the timespan or timestamp. An example is shown in Figure 1.

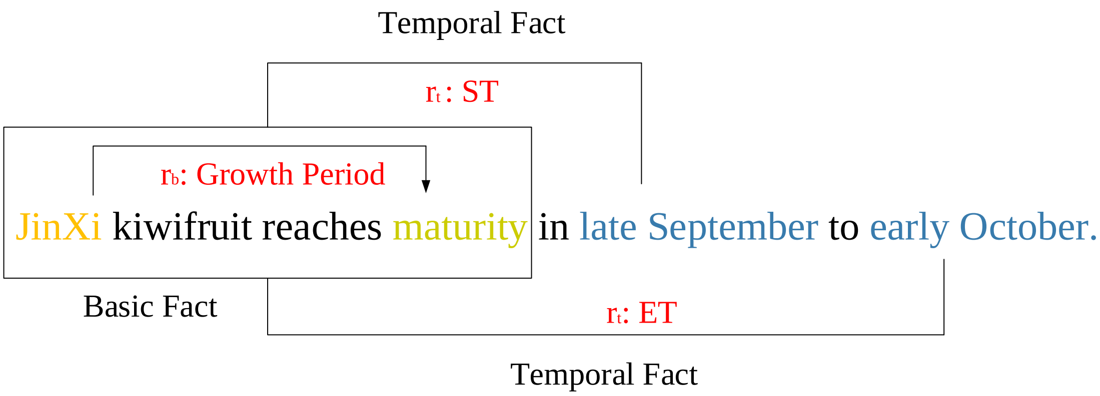

Figure 1. Examples of temporal facts.

3.2. Model design

This paper introduces a joint model to extract temporal instances, which is divided into two sub-modules. The basic fact (s, r, o) is extracted from primary module. Then it is used as the input of the high-level model to determine its corresponding time expression under predefined relations. The framework of the proposed model is shown in Figure 2.

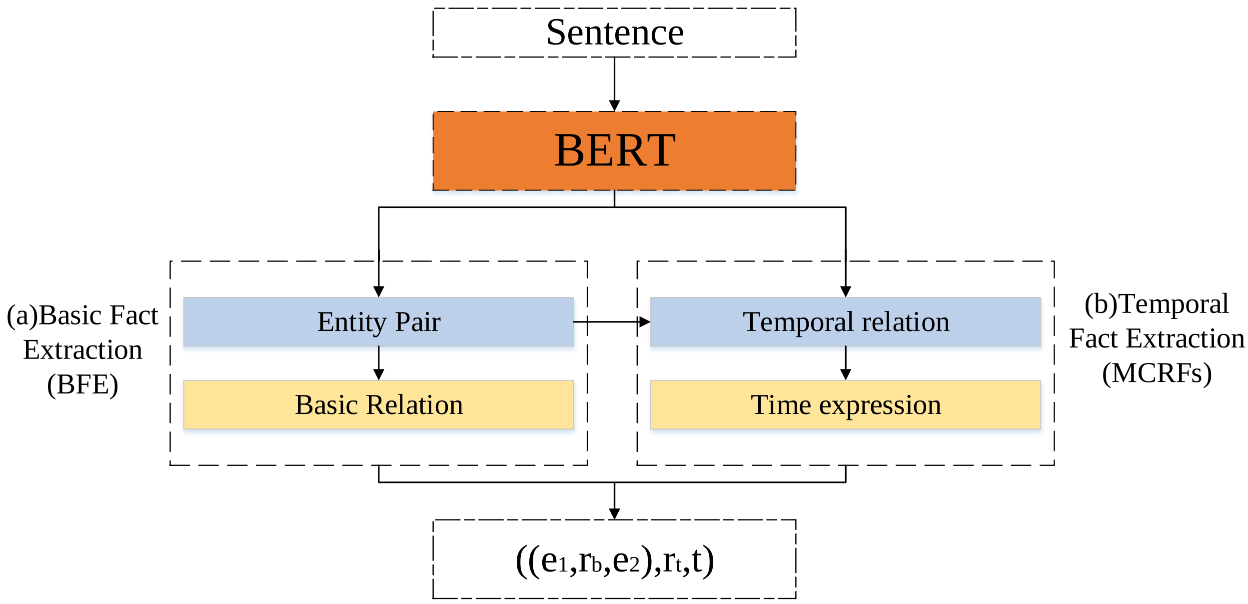

Figure 2. Framework of BFE-MCRFs.

To illustrate the inputs and outputs of each module, we show the overall module in detail in Figure 3.

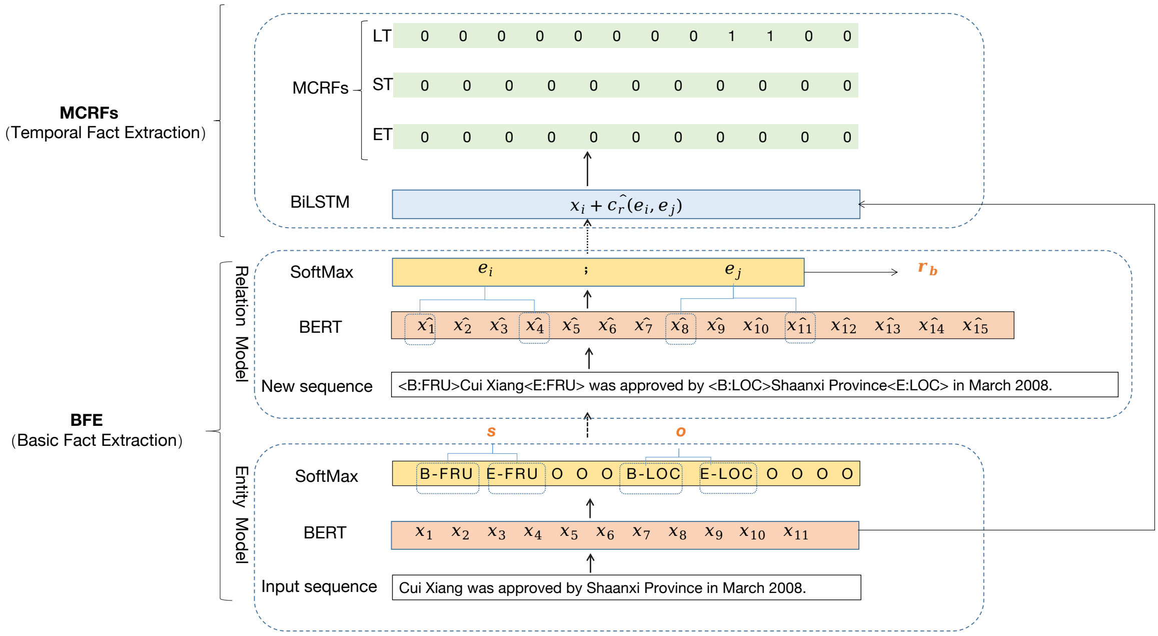

Figure 3. Overall architecture of the proposed framework (BFE-MCRFs).

3.2.1. BERT encoder

The encoder layer extracts feature information from sentences, which is fed into subsequent modules. A pre-trained BERT (Bidirectional Encoder Representation from Transformers) model is employed to encode the contextual information [30]. BERT, a multi-layer bidirectional Transformer based language representation model, is engineered to learn deep semantic representations by the sum of three embeddings: Token Embeddings, Position Embeddings and Segment Embeddings. And it has recently been proven remarkably effective in many downstream tasks [28]. This paper follows this trend and applies the BERT as the encoder.

3.2.2. Basic fact extraction model

The primary module is a conventional entity-relation extraction model, which is applied to extract basic facts from the text. It includes two sub-tasks, named entity recognition and relation classification. We run the entity model to find all possible entities in the sentence, and then for each pair of candidate entities, we apply the relation model to find all relevant relations. The framework of BFE (Basic Fact extraction) is shown in the bottom part of Figure 3.

Entity model

The input of the model is a sentence S consisting of n tokens $s_1, s_2, ..., s_n$.

Let ε and l denote a set of pre-defined entity types and label types, respectively. $\varepsilon = \{type_1, type_2, ..., type_j\}$, $l = \{B - type_j, I - type_j, E - type_j, O\}$. j represents the number of entity type. B, I, E represent the begin, inside and end character of entity, respectively. O represents the non-entity characters. The named entity recognition task is, for each token s_i, to predict a label type $y_e(s_i) \in l$.

The output of the entity model is $Y_e = \{(s_i, l), s_i \in S, l \in \varepsilon\}$. We first use a pre-trained language model (BERT) to obtain contextualized representations $X = x_1, x_2, ..., x_n$, for each input sentence S, where x_i is the character vector of the i_{th} word. Then x_i is fed into the SoftMax layer to predict its label type $l_i \in l$.

$$P_e(l_i|x_i) = softmax(W_e(x_i)) \tag{3.1}$$

Where W_e represents the trainable weight.

The most likely label for each character is calculated by SoftMax, and the final sequence of maximum probability labels for the sentence is formed. Based on the pair of B-$type_j$ and E-$type_j$, all entities contained in the sentence and their types can be identified.

Relation model

Let R denote a set of predefined relation types. The task is, for every pair of candidate entities e_i, e_j which are from the entity model, to predict a relation type $y_r(e_i, e_j) \in R$, or there is no relation between them: $y_r(e_i, e_j) = r_o$.

The input of relation model is a updated sentence $\widetilde{S}$. The label-type markers are inserted at the input layer to highlight the entities and their types, which are defined as $\langle B : \varepsilon_i \rangle$,$\langle E : \varepsilon_j \rangle$ and inserted into S before and after the boundary of entities to obtain $\widetilde{S}$. As shown in Figure 3, the input sentence S "Cui Xiang was approved by Shaanxi Province in March 2008." and its label sequence "B-FRU/E-FRU/O/O/O/B-LOC/E-LOC/O/O/O/O", the new sentence $\widetilde{S}$ is presented as "$\langle B : FRU \rangle$ Cui Xiang $\langle E : FRU \rangle$ was approved by $\langle B : LOC \rangle$ Shaanxi Province $\langle E : LOC \rangle$ in March 2008."

Then $\widetilde{S}$ is encoded by BERT and its contextual representation is $\hat{x}$. The candidate entity pair (e_i, e_j) representation is calculated as follow:

$$c_r(e_i, e_j) = [x_{\hat{start(i)}} + x_{\hat{end(i)}}; x_{\hat{start(j)}} + x_{\hat{end(j)}}] \tag{3.2}$$

Where $start(i)$ and $end(i)$ are the entity indices of $\langle B : \varepsilon_i \rangle$ and $\langle E : \varepsilon_i \rangle$ in $\hat{X}$. Finally, $c_r(e_i, e_j)$ is used to predict the relation type:

$$P_e(r|e_i, e_j) = softmax(W_r(c_r(e_i, e_j))) \tag{3.3}$$

3.2.3. Temporal fact extraction model

The high-level model is designed to extract the temporal information, which is highly related to the basic facts. It has three layers: an input layer, an encoding layer and a sequence tagger. The framework of Multi-layer CRFs (MCRFs) is shown in the upper part of Figure 3.

Let T denote the collection of predefined temporal relations, which includes ST(start time), ET(end time), LT(lasting time) and NT(null time). For a basic triplet $t(e_m, e_n, r)$, the task of this module is to predict its temporal expression, as shown in Eq 3.4. The output of the task is Y_r=((e_m,e_n,r),T),$r \in R$.

$$y_t(ST|t) = t_s, y_t(ET|t) = t_e, y_t(LT|t) = t_l, y_t(NT|t) = null \tag{3.4}$$

Input layers The base input of the high-level module is X. Since FCTD has the feature that the temporal information is highly related to the subject and object, we utilize the entity pair (e_m, e_n) of the basic triplet exported in the primary module as the head entities for temporal fact extraction.

BiLSTM layer In this section, the representation of entity pairs $c_r(e_m, e_n)$ is summed and averaged rather than concatenated in order to keep the input vector dimensionality consistent. The sentence representation X and the embedding of entity pair $\hat{c}_r(e_m, e_n)$ are the input of the BiLSTM. The input

$\hat{X} = \{\hat{x}_1, \hat{x}_2, ..., \hat{x}_n\}$ is calculated as shown in Eq 3.5. Emission score matrix P could obtain from Equation 3.6.

$$\hat{x}_i = x_i + \hat{c}_r(e_m, e_n) \tag{3.5}$$

$$P = BiLSTM(\hat{X}) \tag{3.6}$$

The sequence tagger The transition scores between each label could be learned by CRF. Multi-layer CRFs are applied to output the label sequence of the sentence, which consists of three CRF layers corresponding to three temporal relations. The character position of the time expression is marked as 1. For each relation in T, we recalculate the emission score matrix P as shown in Eq 3.7, Where $r \in (ST, ET, LT)$ represents the relation of the temporal facts, i is the number of relations, W, b are the weight matrix.

$$\hat{P} = W_i^r P + b_i^r \tag{3.7}$$

Then the new emission matrix $\hat{P}$ is input into each CRF to obtain the label sequence to target relation:

$$S(X, y) = \sum_{i=0}^{n} A_{y_i, y_i+1} + \sum_{i=0}^{n} \hat{P}_{i, y_i} \tag{3.8}$$

$$y^* = argmaxS(X, y) \tag{3.9}$$

where i is the number of label tags, and A is the transfer matrix, A_{y_i, y_i+1} denotes the transfer score from the y_i tag to the y_{i+1} tag and y^* denotes the sequence with the highest probability value calculated.

3.3. Loss function

We employ the cross-entropy loss function to define the loss of BFE-MCRFs, which is denoted as l. A basic fact is considered correct only if its entity pairs and the relationships between them are correctly extracted. For each temporal fact, it is considered correct only if the boundaries and contents of the correct temporal expression are extracted under a specific temporal relation.

$$l = \sum_{x=ST,ET,LT} \sum_{k=1}^{m} \sum_{i,j=1,i \neq j}^{n} -logP(t_x | e_i, r_k, e_j) \tag{3.10}$$

3.4. Dataset

To evaluate the effectiveness of the proposed model, we conduct several experiments on two datasets: FCTD and DUIE 2.0. And both datasets have the problem of single entity overlapping. The characteristics and type distribution of each dataset are shown in Figure 4.

FCTD It is a self-constructed dataset on fruit cultivation techniques. These fruits ripened annually. The FCTD is constructed by a Python crawler that crawls fruit planting information from relevant official websites. By analyzing the corpus, we predefine six entity classes: fruit (FRU), cultivation technology (CTEC), function (FUN), period (PER), color(COL) and location (LOC), and nine relation categories: Flesh color, Usage, Pre-growth, Mid-growth, Later-growth, Transplantation process, After, Origin, and Process location. FCTD contains 12381 basic facts and 9409 temporal facts.

DUIE 2.0 It is the largest open Chinese dataset for the complex relationship extraction task. And it could download from https://www.luge.ai/#/luge/dataDetail?id=5. We distill the temporal facts from the whole dataset for the experiment.

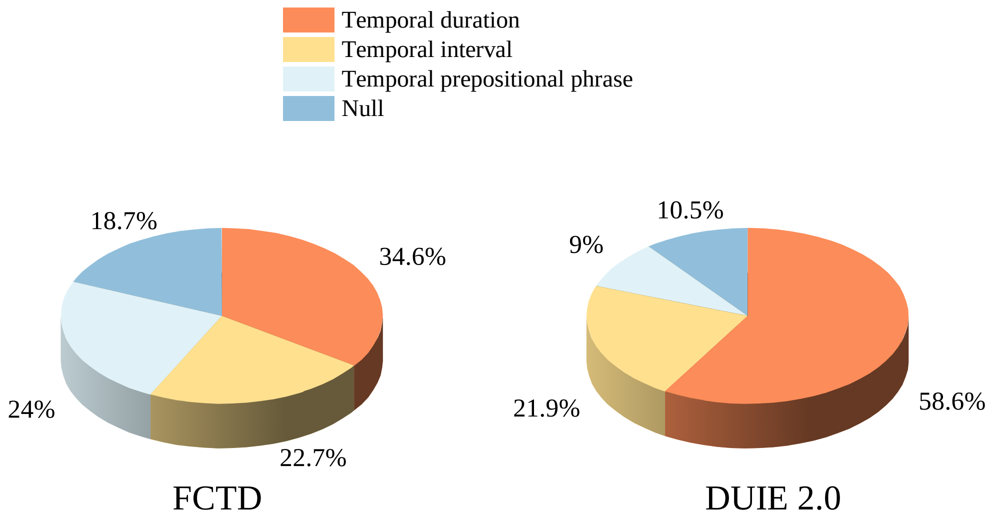

Figure 4. Temporal types distribution of FCTD and DUIE 2.0.

4. Results

4.1. Experimental setup

Implementation We construct our model using TensorFlow with the NVIDIA RTX3060 GPU.

BERT We apply the BERT (pretrained model in Chinese) as the basic encoder for the experiment.

Hyperparameters The batch size is chosen in [64,126,258]. We also adopt 50 training epochs with the early stopping mechanism to prevent overfitting of the model and save the optimal model when the performance does not show any improvement. The parameter settings of BERT and BiLSTM are listed in Table 2. We use the Adam (Adaptive momentum) optimizer [31] and fix the hyperparameters on the validation sets.

Evaluation Metrics In this work, we propose a joint model BFE-MCRFs to extract temporal facts from free text. Hence, we first measure the quality of basic facts and then evaluate the quality of temporal facts. To verify the validity of the BFE-MCRFs, We follow the standard evaluation protocol and employ precision P, recall R, F1 score as primary evaluation metrics. We split dataset into 70% for training to design our model, 10% for validation to tune parameters and 20% for testing to decided on our final model. The fixed length of a sentence is set at 100 as previous works suggest [32].

Table 2. The parameter settings of BERT and BiLSTM.

Model	Parameters
BERT	12 layers 768 dimensions learning rate 5e-5 pad size 128 Tanh function
BiLSTM	2 layers 128 dimensions learning rate 1e-3 pad size 128 ReLu function Tanh function

4.2. Experimental Results

Compared Methods To verify the effectiveness of the proposed model, we compare it with the classical neural model for entity and relation extraction, including BERT-CRF [33], BiLSTM-CRF [34], BERT-BiLSTM-CRF [27].

Main Results Table 3 shows the results of different baselines for temporal fact extraction on two datasets. The BFE-MCRFs model overwhelmingly outperforms all the baselines in terms of all three-evaluation metrics, achieving 95.2 and 92.7% in terms of the F1 score on FCTD and DUIE 2.0 datasets, respectively. Moreover, compared with BERT-BiLSTM-CRF, our model achieves a significant increase in F1. Specifically, BFE-MCRFs achieves improvements of 12.5 and 11.4% in the F1 score, respectively. This demonstrates the effectiveness of the proposed model.

Table 3. Comparison of experimental results

	FCTD			DUIE 2.0		
Model	P	R	F1	P	R	F1
BiLSTM-CRF	0.678	0.701	0.689	0.665	0.689	0.677
BERT-CRF	0.77	0.732	0.751	0.734	0.715	0.724
BERT-BiLSTM-CRF	0.81	0.845	0.827	0.806	0.821	0.813
BFE-MCRFs	**0.956**	**0.948**	**0.952**	**0.943**	**0.911**	**0.927**

As shown in Table 3, BERT-CRF performs better than BiLSTM-CRF on both datasets because BERT can capture better contextual features than BiLSTM. Among the three baseline models, BERT-BiLSTM-CRF has the highest F1 value, and in many studies, BERT-BiLSTM-CRF has proved to be very effective in the entity relationship extraction task. In extracting temporal facts ((s, r, o), t), the classical model first extracts all entities (s, o, t) and then classifies their relations to obtain r. However, stepwise extraction leads to some basic facts not matching their temporal expressions. BFE-MCRFs applies a joint model to extract temporal facts. And it adds temporal relationships between the basic facts and the temporal expressions. It makes use of the basic facts as head entities to extract the temporal expressions under the corresponding temporal relations.

It can also be observed from Figure 5 that there are obvious differences in the performance of the model on FCTD and DUIE 2.0 datasets. The analysis reveals it is caused by the different distribution of the data. More precisely, as indicated in Figure 4, the FCTD dataset consists all types of temporal expression, whereas most sentences in the DUIE 2.0 dataset belong to temporal duration extraction. This inconsistent data distribution of the two datasets leads to comparatively better performance on FCTD and worse performance on DUIE 2.0 for all the evaluation metrics, which reflects the difficulty of the task in extracting complex temporal facts.

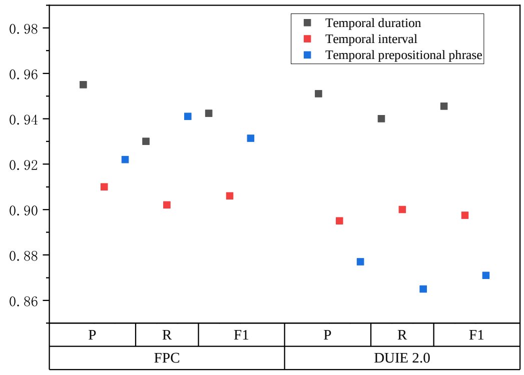

Figure 5. The performance of BFE-MCRFs on different temporal relations.

To further investigate the effects of BFE-MCRFs in terms of basic fact extraction and single entity overlapping extraction, ablation experiments on the FCTD and DUIE 2.0 datasets are constructed.

As shown in Table 4, both parts can assist BFE-MCRFs in jointly extracting temporal facts efficiently, where BFE seems to play a more significant role. When we replace the BFE with a classical extraction model, the performance decreases by 26.3 and 24.7% in terms of the F1 score, respectively. The main reasons are that BFE can extract more correct basic facts to improve the accuracy of extraction of high-level model, reducing the impact of noise during the whole extraction process, and provide better support for temporal extraction. Furthermore, upon removing MCRFs, the performance decreases by 9.2% and 7.8% in terms of the F1 score, respectively. The analysis of the extraction results illustrates that MCRFs can naturally capture the single entity overlapping relations in sentences.

Table 4. Ablation study of BFE-MCRFs on the FCTD and DUIE 2.0 datasets.

	FCTD			DUIE 2.0		
Model	P	R	F1	P	R	F1
BFE-MCRFs	**0.956**	**0.948**	**0.952**	**0.943**	**0.911**	**0.927**
Replace MCRFs	0.854	0.867	0.86	0.843	0.855	0.849
Replace BFE	0.678	0.701	0.689	0.669	0.692	0.68

5. Discussion

To illustrate the problem, we list a few examples (the origin sentence is in Chinese) from FCTD in Table 5. As can be seen, BFE-MCRFs could identify one or more temporal facts in each sentence. This demonstrates the effectiveness of the proposed model in solving the problem of extracting temporal facts in each case.

Table 5. Case study of BFE-MCRFs and baseline model on the FCTD dataset.

Sentence1	Spring planting of luochuan apples takes place in late March to early April.
BERT+BiLSTM+CRF	((luochuan, pre-growth, spring planting), NULL)
	((NULL, pre-growth, spring planting), late March)
	((NULL, pre-growth, spring planting), early April)
	((luochuan, pre-growth, NULL), late March)
	((luochuan, pre-growth, NULL), early April)
Our model	((luochuan, pre-growth, spring planting), ST, late March)
	((luochuan, pre-growth, spring planting), ET, early April)
Ground truth	((luochuan, pre-growth, spring planting), (late March, early April))
Sentence2	Hong Yang can be stored for 7 days at room temperature.
BERT+BiLSTM+CRF	((Hong Yang, late growth, store), NULL)
	((NULL, late growth, store), 7 days)
Our model	((Hong Yang, late growth, store), LT, 7 days))
Ground truth	((Hong Yang, late growth, store), 7 days))
Sentence3	Yate can be stored for 3 days after ripening.
BERT+BiLSTM+CRF	((ripen, after, store), NULL)
	((NULL, after, store), 3 days)
Our model	((ripen, after, store), LT, 3 days))
Ground truth	((ripen, after, store), 3 days))

As indicated in Table 5, the first example contains a temporal interval and relations (ST and ET). The second example includes temporal duration and relation (LT). The third example contains temporal prepositional phrase and relation (LT). For the testing of the BERT-BiLSTM-CRF model, the following can be observed: (1) it could identify the basic facts precisely; (2) even though it could extract the time expression, it regards temporal fact extraction as simple fact extraction. It could not extract $((s, r, o), t)$ in a single model.

6. Conclusions

This paper proposes BFE-MCRFs, an end-to-end neural network model for extracting all temporal facts in text corpus. We input basic facts to the high-level model in order to extract temporal expression and define the link between basic facts and time expressions as relations. Experiments on the FCTD and DUIE 2.0 datasets indicate that BFE-MCRFs outperforms BERT+BiLSTM+CRF by 12.5 and 11.3%. Moreover, the proposed model achieves state-of-the-art performance in terms of temporal fact extraction using the neural network model. The proposed work has considerable potential for further applications. The proposed model BFE-MCRFs designs a general framework to extract temporal facts from free text, which could be used for various tasks in information extraction.

The key components in the above general framework can be instant in many ways to improve the performance of the model, such as accuracy, computing speed or generalization ability.

In the future work, First, for the purpose of increasing model performance improvement and test credibility, the dataset of time object-oriented extraction method research should be further expanded in depth and breadth. In this regard, we will explore cooperation with researchers with common research interests. Second, the current treatment of entity pairs in the model is a compromise on the input dimension, but this treatment does pose a potential threat to the robustness of the model. Therefore, it is necessary to further study better treatment methods to balance these two points. Third, if those new models derived from the original BERT model were used to code after adaptation processing, whether these models can play their advantages in this model without introducing other defects needs further research.

Acknowledgments

This work was supported by Youth Scholars of Beijing Technology and Business University [grant numbers 19008022169] and the National Key Research and Development Program of China [grant numbers 2021YFD2100605] and Beijing Science and Technology Planning Project [grant numbers Z221100007122003].

Conflict of interest

The authors declare there is no conflict of interest.

References

1. J. Yan, C. Wang, W. Cheng, M. Gao, A. Zhou, A retrospective of knowledge graphs, *Front. Comput. Sci.*, **12** (2018), 55–74. https://doi.org/10.1007/s11704-016-5228-9 .
2. T. Mitchell, W. Cohen, E. Hruschka, P. Talukdar, B. Yang, J. Betteridge, et al., Never-ending learning, *Commun. ACM*. **61** (2018), 103–115. https://doi.org/10.1145/3191513
3. W. Wu, H. Li, H. Wang, K. Q. Zhu, Probase: a probabilistic taxonomy for text understanding, in *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, 2012, 481–492. https://doi.org/10.1145/2213836.2213891
4. T. Rebele, F. Suchanek, J. Hoffart, J. Biega, E. Kuzey, G. Weikum, YAGO: A multilingual knowledge base from wikipedia, wordnet, and geonames, in *The Semantic Web – ISWC 2016* , (2016), 177–185. https://doi.org/10.1007/978-3-319-46547-0_19
5. I. Mani, Recent developments in temporal information extraction, in *Recent Advances in Natural Language Processing III*, 2003. https://doi.org/10.1075/cilt.260.06man
6. C. Lim, Y. Jeong, H. Choi, Survey of temporal information extraction, *J. Inf. Process. Sys.*, **15** (2019), 931–956.
7. Y. Cao, W. Groves, T. K. Saha, J. Tetreault, A. Jaimes, H. Peng, et al., XLTime: A cross-lingual knowledge transfer framework for temporal expression extraction, in *Findings*

of the Association for Computational Linguistics: NAACL 2022, (2022), 1931–1942. http://doi.org/10.18653/v1/2022.findings-naacl.148

8. X. Ling, D. S. Weld, Temporal information extraction, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, (2010), 1385–1390.

9. H. Li, J. Strötgen, J. Zell, M. Gertz, Chinese temporal tagging with heidelTime, in *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, volume 2: Short Papers*, (2014), 133–137. http://doi.org/10.3115/v1/E14-4026

10. W. Li, K. Wong, C. Yuan, Toward automatic Chinese temporal information extraction, *J. Am. Soc. Inf. Sci. Technol.*, **52** (2001), 748–762. https://doi.org/10.1002/asi.1126

11. H. Tanev, J. Piskorski, M. Atkinson, Real-time news event extraction for global crisis monitoring, in *Natural Language and Information Systems*, 207–218, https://doi.org/10.1007/978-3-540-69858-6_21

12. J. Strötgen, M. Gertz , P. Popov, Extraction and exploration of spatio-temporal information in documents, in *Proceedings of the 6th Workshop on Geographic Information Retrieval*, (2010), 1–8. https://doi.org/10.1145/1722080.1722101

13. N. Kannen, U. Sharma, S. Neelam, D. Khandelwal, S. Ikbal, H. Karanam, et al., Targeted extraction of temporal facts from textual resources for improved temporal question answering over knowledge bases, preprint, arXiv:2203.11054.

14. I. Mani, B. Schiffman, J. Zhang, Inferring temporal ordering of events in news, in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology: companion volume of the Proceedings of HLT-NAACL 2003*, **2** (2003), 55–57. https://doi.org/10.3115/1073483.1073502

15. J. Pustejovsky, P. Hanks, R. Saurí, A. See, R. Gaizauskas, A. Setzer, et al., The TIMEBANK Corpus, *Nat. Lang. Process. Inf. Syst.*, **4592** (2002), 647–656,

16. Y. Wang, M. Zhu, L. Qu, M. Spaniol, G. Weikum, Timely YAGO: harvesting, querying, and visualizing temporal knowledge from Wikipedia, in *Proceedings of the 13th International Conference on Extending Database Technology*, (2010), 697–700, https://doi.org/10.1145/1739041.1739130

17. D. Vrandečić, M. Krötzsch, Wikidata: A free collaborative knowledgebase, *Commun. ACM*, **57** (2014), 78–85, https://dl.acm.org/doi/10.1145/2629489

18. J. Hoffart, F. M. Suchanek, K. Berberich, G. Weikum, YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia, *Artif. Intell.*, **194** (2013), 28–61. https://doi.org/10.1016/j.artint.2012.06.001

19. E. Kuzey, G. Weikum, Extraction of temporal facts and events from Wikipedia, in *Proceedings of the 2nd Temporal Web Analytics Workshop*, (2012), 25–32. https://doi.org/10.1145/2169095.2169101

20. Y. Liu, W. Hua, X. Zhou, Temporal knowledge extraction from large-scale text corpus, *World Wide Web*, **24** (2021), 135–156, https://doi.org/10.1007/s11280-020-00836-5

21. B. Tang, Y. Wu, M. Jiang, Y. Chen, J. C. Denny, H. Xu, A hybrid system for temporal information extraction from clinical text, *J. Am. Med. Inform. Assoc.*, **20** (2013), 828–835. https://doi.org/10.1136/amiajnl-2013-001635
22. G. Moharasan, T. B. Ho, Extraction of temporal information from clinical narratives, *J. Healthc. Inform. Res.*, **3** (2019), 220–244. https://doi.org/10.1007/s41666-019-00049-0
23. R. Han, Q. Ning, N. Peng, Joint event and temporal relation extraction with shared representations and structured prediction, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, (2019), 434–444. http://doi.org/10.18653/v1/D19-1041
24. R. Han, Y. Zhou, N. Peng, Domain knowledge empowered structured neural net for end-to-end event temporal relation extraction, in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (2020), 5717–5729. http://doi.org/10.18653/v1/2020.emnlp-main.461
25. C. Lin, T. Miller, D. Dligach, S. Bethard, G. Savova, Representations of time expressions for temporal relation extraction with convolutional neural networks, in *BioNLP 2017*, (2017), 322–327. http://doi.org/10.18653/v1/W17-2341
26. P. Cao, X. Zuo, Y. Chen, K. Liu, J. Zhao, W. Bi, Uncertainty-aware self-training for semi-supervised event temporal relation extraction, in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, (2021), 2900–2904. https://doi.org/10.1145/3459637.3482207
27. H. Wen, H. Ji, Utilizing relative event time to enhance event-event temporal relation extraction, in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, (2021), 10431–10437. http://doi.org/10.18653/v1/2021.emnlp-main.815
28. K. Ma, Extraction of temporal information from social media messages using the BERT model, *Earth. Sci. Inform.*, **15** (2022), 573–584. https://doi.org/10.1007/s12145-021-00756-6
29. A. Uzun, A. C. Tantuğ, ITUTime: Turkish temporal expression extraction and normalization, in *Distributed Computing and Artificial Intelligence*, **2** (2021), 74–85. https://doi.org/10.1007/978-3-030-86887-1_7
30. J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (2019), 4171–4186. http://doi.org/10.18653/v1/N19-1423
31. D. Kingma, J. Ba, Adam: A method for stochastic optimization, preprint, arXiv:1412.6980.
32. T. Fu, P. Li, W. Ma, GraphRel: Modeling text as relational graphs for joint entity and relation extraction, in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (2019), 1409–1418. http://doi.org/10.18653/v1/P19-1136
33. L. Mingyi, T. Zhiying, Z. Tong, S. Tonghua, X. Xiaofei, Z. Wang, Ltp: A new active learning strategy for crf-based named entity recognition, *Neural Process. Lett.*, **54** (2022), 2433–2454. https://doi.org/10.1007/s11063-021-10737-x

34. N. Deng, F. Hao, C. Xu, Named entity recognition of traditional chinese medicine patents based on bilstm-crf, *Wireless Commun. Mobile Comput.*, **2021** (2021), 6696205. https://doi.org/10.1155/2021/6696205