Joint Modeling for Query Expansion and Information Extraction with Reinforcement Learning

Motoki Taniguchi, Yasuhide Miura and **Tomoko Ohkuma**

Fuji Xerox Co., Ltd.

{motoki.taniguchi, yasuhide.miura, ohkuma.tomoko}@fujixerox.co.jp

Abstract

Information extraction about an event can be improved by incorporating external evidence. In this study, we propose a joint model for pseudo-relevance feedback based query expansion and information extraction with reinforcement learning. Our model generates an event-specific query to effectively retrieve documents relevant to the event. We demonstrate that our model is comparable or has better performance than the previous model in two publicly available datasets. Furthermore, we analyzed the influences of the retrieval effectiveness in our model on the extraction performance.

1 Introduction

Information extraction about an event is gaining growing interest because of the increases in text data. The task of information extraction about an event from a text is defined to identify a set of values including entities, temporal expressions and numerical expressions that serve as a participant or attribute of the event. The extracted information is useful for various applications such as risk monitoring (Borsje et al., 2010) and decision making support (Wei and Lee, 2004).

Conventional information extraction systems provide higher performance if the amount of labeled data is larger. Labeled training data is expensive to produce and thus the data amount is limited. In this case, extraction accuracy can be improved using an alternative approach that incorporates evidence from external sources such as the Web (Kanani and McCallum, 2012; Narasimhan et al., 2016; Sharma et al., 2017). However, this approach faces the following challenges: issuing an effective query to the external source, identifying documents relevant to a target event from retrieval results and reconciling the values extracted from the relevant documents.

To overcome these problems, several attempts have been made to model the decisions as a Markov Decision Process (MDP) with deep reinforcement learning (RL) (Narasimhan et al., 2016; Sharma et al., 2017). The agent of these models is trained to maximize expected rewards (extraction accuracy) by performing actions to select an expanded query for external source and to reconcile values extracted from documents retrieved from the external source. The models use the title of source document as an original query and templates to expand this query. Expansion terms of the template are same in any event even though an optimal query depends on the event. Therefore, it is still a challenge to issue an effective query to an external source.

In this study, we extended the previous models by introducing a query expansion based on pseudo-relevance feedback (PRF) (Xu and Croft, 1996). The PRF based query expansion assumes that the top-ranked documents retrieved by an original query are relevant to the original query. An agent of our model selects a term from those documents and generates an expanded query by adding the term into the original query. The PRF based query expansion enables us to add an event-specific term into the query without additional resources. For instance, let us consider an information extraction about a shooting incident. The query "Shooting on Warren Ave. leaves one dead" retrieves the documents which may contain the term "New York". The addition of the event-specific term "New York" to the query leads to the filtering out of irrelevant documents, and thus to improves retrieval performance because "Warren Ave." is located in "New York". Therefore, we expect to improve extraction accuracy by introducing PRF based query expansion.

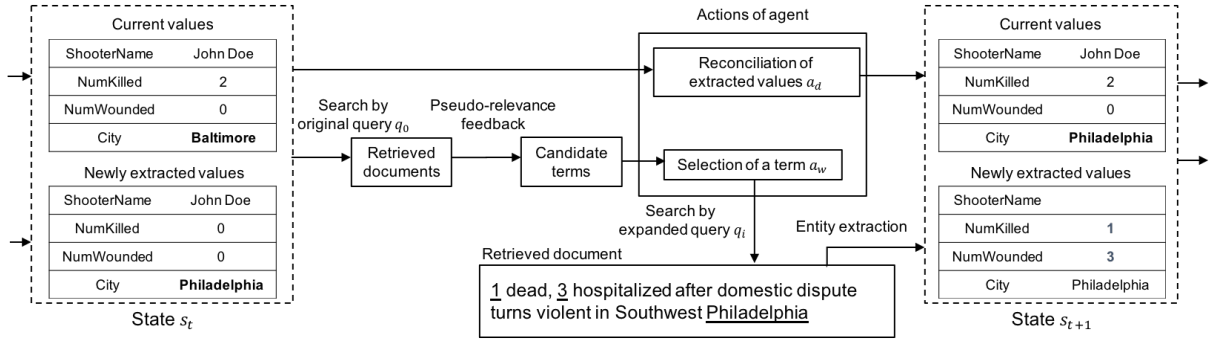In contrast to the previous models, candidate terms for query expansion in our model vary de-

Figure 1: Illustration of states, actions and state transition in our framework. For simplicity, only the sets of current and newly extracted values are shown for the states.

pending on an event. Therefore, we exploit an original query and its candidate terms information as inputs of policy networks in addition to state information.

The contributions of this paper are follows:

- We propose a joint model for PRF based query expansion and information extraction with RL.

- We investigated the oracle extraction accuracy as an indicator of the model's retrieval performance to reveal that the PRF based query expansion outperforms the template query in two publicly available datasets.

- We demonstrate that our model is comparable or better extraction performance compared to the previous model in the datasets.

2   Related Work

Information extraction incorporating external sources: Information extraction incorporating external sources has been increasingly investigated in knowledge base population (Ji and Grishman, 2011; West et al., 2014) and multiple document extraction (Mann and Yarowsky, 2005). In contrast to the tasks of these studies, a challenge exists in our task to identify documents relevant to a target event, and this complicates extraction of information.

Narasimhan et al. (2016) and Sharma et al. (2017) modeled the information extraction tasks as an MDP with RL. They demonstrated that their models outperformed conventional extractors and meta-classifier models. There are two crucial differences between our model and aforementioned models. First, the proposed model is trained to select an optimal term of query expansion for each

original query instead of a query template. Second, the proposed model also leverages an original query and its candidate terms as the input of policy networks, whereas the above-mentioned models use only state information.

Query expansion: Query expansion can be categorized into global and local methods (Manning et al., 2008). The global methods include query expansion using a manual thesaurus (Voorhees, 1994), and an automatically generated thesaurus based on word co-occurrence statistics over corpus (Qiu and Frei, 1993). The template query that is used in the previous RL based model belongs to the global methods.

There are several approaches of local methods that use query log (Cui et al., 2002), and are based on PRF (Xu and Croft, 1996). We employed a PRF based method because it does not require additional resources. Moreover, local methods have been evidenced to outperform global methods in information retrieval (Xu and Croft, 1996).

Nogueira and Cho (2017) proposed an RL based approach to model query reformulation in information retrieval. They also reformulated a query through PRF. In contrast, our proposed approach targets information extraction rather than document retrieval. The goal of our task is to extract multiple values of event attributes as well as to retrieve relevant documents. Moreover, the document collection in the RL-based approach (Nogueira and Cho, 2017) is limited to Wikipedia pages or academic papers, while we use the Web as an open document-collection platform.

3   Proposed Model

3.1   Framework

We model the information extraction task as an MDP in a similar manner to that by

Narasimhan et al. (2016) and Sharma et al. (2017); however, the query-selection strategy differs. An agent in our framework selects a term to add to an original query instead of selecting a query template. In this section, we mainly describe the difference between our framework and the previous framework.

At the beginning of each episode, an agent is given a source document to extract information about an event. Figure 1 illustrates an example of state transition. The state comprises the confidence score of current and newly extracted values, the match between current and new values, term frequency-inverse document frequency (TF-IDF) of context words, and TF-IDF similarity between source and current documents, similar to the method by Narasimhan et al. (2016). At each step, the agent performs two actions: a reconciliation decision a_d, which involves the accepting of extracted values for one or all attributes or rejecting all newly extracted values (or ending an episode), and a term selection for query expansion a_w. In the example, the agent takes a reconciliation decision a_d to update the value of *City* attribute from *Baltimore* to *Philadelphia* based on the state s_t in time step t. Simultaneously, the agent performs term selection a_w that is used to form the expanded query to retrieve the next document. The candidate terms $W = \{w_1, w_2, \cdots, w_N\}$ are collected from the documents retrieved by an original query q_0. Here, we use the title of the source document as q_0. The agent selects a term w_i from W and generates expanded query q_i by adding w_i into q_0. q_i is used to retrieve the next document and the new values are extracted from the document by base extractor. State s_{t+1} of the next step is determined according to the updated values of the event attributes and the newly extracted values. The agent receives a reward, which is defined as the difference between the accuracy in the current time step and the previous time step. We add a negative constant to the reward to avoid long episodes. In the subsequent time steps, the agent sequentially chooses the two actions a_d and a_w until a_d is a stop decision. For more details of state, reward and base extractor, refer to (Narasimhan et al., 2016).

The agent is trained to optimize the expected rewards by choosing actions to select an expansion term for query expansion and to reconcile values extracted from documents.
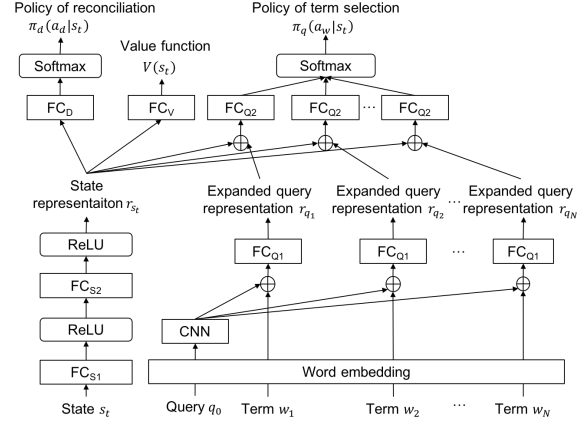

Figure 2: Overview of our model. FC: fully connected layer; $\oplus$: concatenation of vectors.

3.2 Network Architecture

We use neural networks to model decision policy $\pi_d(a_d|s_t)$ and term selection policy $\pi_q(a_w|s_t)$ as probability distributions over candidate actions a_d and a_w. Figure 2 represents an overview of our policy networks. Decision policy $\pi_d(a_d|s_t)$ and value function $V(s_t)$ is calculated using a state representation r_{s_t} that is obtained with two fully connected layers in the same manner as in Sharma et al. (2017).

In contrast to the previous framework, candidate terms in our framework depend on an event. Hence, we utilize a pairwise interaction function whose input is the state representation r_{s_t} and expanded query representation r_{q_i} to calculate term selection policy $\pi_q(a_w|s_t)$. The words of q_0 are embedded using a word embedding layer and process using a convolutional neural network (CNN) and a max pooling layer, similar to the method by Kim (2014). r_{q_i} for each term w_i is obtained by feeding the concatenation of the output of the max pooling layer and the word embedding of a candidate term w_i to a fully connected layer FC_{Q1}. We feed the concatenation of r_{s_t} and r_{q_i} for each term w_i to a fully connected layer FC_{Q2}. The parameters of the FC_{Q1} and FC_{Q2} are shared among the candidate terms. $\pi_q(a_w|s_t)$ is obtained by normalizing the outputs of the FC_{Q2} over the candidate terms.

We train the policy and value networks by using the Asynchronous Advantage Actor-Critic (A3C) (Mnih et al., 2016) as used in Sharma et al. (2017). A3C can speed up the learning process of the policy networks by training multiple agents asynchronously. Further details on the A3C can be found in Sharma et al. (2017).

Documents		Shooting			Adulteration		
		train	dev	test	train	dev	test
Source document		372	146	146	292	42	148
Template query	N=5/4	22,697	8,874	9,039	11,425	1,597	6,166
PRF based expanded query	N=5/4	20,243	8,877	8,128	10,994	1,383	5,460
	N=11	45,432	19,505	18,372	30,154	3,965	14,847
	N=21	90,006	37,256	36,013	57,636	7,785	27,963
	N=31	134,475	54,966	54,021	84,768	11,638	41,181

Table 1: The number of source documents and downloaded documents in each set.

4 Experiments

4.1 Experimental Setting

We evaluated our model on *Shooting* and *Adulteration* datasets that used in Narasimhan et al. (2016). For each an original query, we collected the candidate terms obtained from the first M words of the top-K documents retrieved through Bing Search API[1]. The vocabulary size of the candidate terms N is defined as $MK + 1$ because the null token, namely no query expansion, is also included in the candidate terms. We downloaded the top 20 documents from the Bing Search API as the external sources through an expanded query. Statistics of the original datasets and downloaded documents is described in Table 1.

Word embeddings are set to fixed vectors of 300 dimensions and is initialized with word2vec embedding trained on Google News Dataset[2]. We set the unit size of FC_{S1} and FC_{S2} to 20, FC_{Q1} to 300, FC_V and FC_{Q2} to 1. Further, we set the number of feature maps of CNN to 200 and the window size of CNN to 3. Discount factor and the constant of entropy regularization were set to 0.8 and 0.01, respectively. We utilize RMSprop (Hinton et al., 2012) as the optimizer and set the number of threads in A3C to 16.

We employed RLIE-A3C (Sharma et al., 2017) as a baseline model to compare with our model. We used their public implementation[3].

4.2 Results

We evaluated the average extraction accuracy for attributes as done in Sharma et al. (2017). Table 2 shows the results of *Shooting* and *Adulteration*.

Our model achieved 1.9 pt increase of average accuracy in *Shooting* and 0.2 pt decrease in *Adulteration* against the RLIE-A3C model when the number of expanded queries N was 5 for *Shooting* and 4 for *Adulteration*; these correspond to $(M, K) = (4, 1)$ or $(3, 1)$ respectively. We varied (M, K) to (10, 1), (10, 2) and (10, 3), which indicate $N = 11, 21$ and 31, to evaluate the effect of the number of expanded queries. The accuracy in our models rarely changes even though the number of expanded queries increases.

Model	External source		Shooting	Adulteration
	Query expansion	# query *N*	Average accuracy	
RLIE-A3C	Template	5(Shooter)/ 4(Adulteration)	0.646	0.612
Our model	PRF		0.665	0.610
		11	0.656	0.611
		21	0.666	0.607
		31	0.653	0.608

Table 2: Extraction accuracy of RLIE-A3C and our models on *Shooting* and *Adulteration* datasets. Underlined values represent the best values on each datasets.

4.3 Discussion

We evaluated oracle extraction accuracy to determine the effectiveness of PRF-based query expansion and to discover why our model does not perform well in *Adulteration*. The oracle extraction accuracy is calculated if an agent perfectly takes actions to select a term for query and reconcile the values from the documents. In other words, the oracle extraction accuracy can be regarded as an indicator of retrieval performance.

Table 3 presents the oracle extraction accuracies from only source documents, the documents retrieved by the original queries and the documents retrieved by the expanded queries by using templates and PRF. Compared with template queries, PRF based query expansion with the same number of queries performed better in *Shooting* and *Adulteration*. Oracle extraction accuracy further im-

Documents	# query *N*	Shooting	Adulteration
Source documents	-	0.612	0.580
Retrieved documents for original queries	1	0.747	0.560
Retrieved documents for template expanded queries	5(Shooting)/ 4(Adulteration)	0.768	0.592
Retrieved documents for PRF based expanded queries (ours)		0.807	0.608
	11	0.820	0.642
	21	0.823	0.663
	31	0.830	0.665

Table 3: Oracle extraction accuracy. "Documents" refers to documents to extract information.

Shooting	1 Dead, 3 Wounded In Shooting On Roxbury Street
	Police: 3 Dead, 1 In Surgery After Dallas Shooting CBS Dallas
	San Francisco police confirm 4 dead following shooting in Hayes Valley
Adulteration	Subway investigates china media reports of doctored expiry dates
	High level of seafood fraud found Denmark
	Fake honey worthing leads to trading standards prosecution

Table 4: Examples of titles of source documents in *Shooting* and *Adulteration* datasets

proved with the increases in the number of queries N. However, no difference was found in its extraction performance (see Table 2). This indicates that increasing the number of queries N complicates the selection of an optimal term for query expansion.

Compared to *Shooting*, the oracle accuracy of the original queries in *Adulteration* is relatively low. Therefore, the assumption that the top-ranked documents retrieved by the original query are relevant to the original query is not satisfied in *Adulteration*. We consider that this is why our model does not achieve an improvement in the extraction performance on *Adulteration*. Table 4 shows examples of the titles of source document used as original queries. We can observe that named entity and numerical expression appeared in more titles of the source documents in *Shooting* than those in *Adulteration*. Therefore, the original queries in *Adulteration* lack specifics and weaken the extraction performance.

5 Conclusions

We integrated the PRF based query expansion to the task of information extraction using RL. Our model can expand a query into an event-specific query without additional resources. To integrate the PRF based query expansion, we introduced a pairwise interaction function to calculate term selection policy $\pi_q(a_w|s)$. Experimental results showed that our model is comparable or better than the previous model in terms of extraction performance in two datasets. Furthermore, we analyzed the relationship between retrieval effectiveness and extraction performance.

In the future work, we plan to develop a model that can generate a complete term sequence of the expanded query rather than adding a term to a query.

References

Jethro Borsje, Frederik Hogenboom, and Flavius Frasincar. 2010. Semi-automatic financial events discovery based on lexico-semantic patterns. *Int. J. Web Eng. Technol.*, 6(2):115–140.

Hang Cui, Ji-Rong Wen, Jian-Yun Nie, and Wei-Ying Ma. 2002. Probabilistic query expansion using query logs. In *Proceedings of the 11th International Conference on World Wide Web*, WWW '02, pages 325–332, New York, NY, USA. ACM.

Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. 2012. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent.

Heng Ji and Ralph Grishman. 2011. Knowledge base population: Successful approaches and challenges. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 1148–1158. Association for Computational Linguistics.

Pallika H. Kanani and Andrew K. McCallum. 2012. Selecting actions for resource-bounded information extraction using reinforcement learning. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, WSDM '12, pages 253–262, New York, NY, USA. ACM.

Yoon Kim. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751. Association for Computational Linguistics.

Gideon Mann and David Yarowsky. 2005. Multifield information extraction and cross-document fusion. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 483–490. Association for Computational Linguistics.

Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press, New York, NY, USA.

Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. 2016. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937.

Karthik Narasimhan, Adam Yala, and Regina Barzilay. 2016. Improving information extraction by acquiring external evidence with reinforcement learning. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages

2355–2365. Association for Computational Linguistics.

Rodrigo Nogueira and Kyunghyun Cho. 2017. Task-oriented query reformulation with reinforcement learning. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 574–583. Association for Computational Linguistics.

Yonggang Qiu and Hans-Peter Frei. 1993. Concept based query expansion. In *Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '93, pages 160–169, New York, NY, USA. ACM.

Aditya Sharma, Zarana Parekh, and Partha Talukdar. 2017. Speeding up reinforcement learning-based information extraction training using asynchronous methods. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2658–2663. Association for Computational Linguistics.

Ellen M. Voorhees. 1994. Query expansion using lexical-semantic relations. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '94, pages 61–69, New York, NY, USA. Springer-Verlag New York, Inc.

Chih-Ping Wei and Yen-Hsien Lee. 2004. Event detection from online news documents for supporting environmental scanning. *Decis. Support Syst.*, 36(4):385–401.

Robert West, Evgeniy Gabrilovich, Kevin Murphy, Shaohua Sun, Rahul Gupta, and Dekang Lin. 2014. Knowledge base completion via search-based question answering. In *WWW*.

Jinxi Xu and W. Bruce Croft. 1996. Query expansion using local and global document analysis. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '96, pages 4–11, New York, NY, USA. ACM.