

少量学習データと Bi-directional LSTM-CNN-CRF による参考文献書誌情報抽出

荒川 瞭平[†] 太田 学^{††} 金澤 輝一^{†††} 高須 淳宏^{†††}

[†] 岡山大学工学部情報系学科 〒700-8530 岡山市北区津島中 3-1-1

^{††} 岡山大学大学院自然科学研究科 〒700-8530 岡山市北区津島中 3-1-1

^{†††} 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: [†]p7zq0av2@s.okayama-u.ac.jp, ^{††}ohta@de.cs.okayama-u.ac.jp, ^{†††}{tkana, takasu}@nii.ac.jp

あらまし 膨大な文書が格納されている電子図書館の運用には、書誌情報データベースの整備が必須である。特に学術論文の参考文献欄には著者名やタイトルなどの有用な書誌情報が集約されているため、参考文献文字列から書誌情報を自動抽出する研究が行われている。浪越らは、参考文献書誌情報抽出において、Bi-directional LSTM-CNN-CRF を利用し、参考文献文字列からトークンとその書誌要素を同時に推定する手法を提案したが、学習データ量がその抽出精度に与える影響は明らかではなかった。そこで本稿では、少量の学習データで浪越らの利用したモデルの書誌情報抽出精度を向上させるため、能動サンプリングとデータ拡張、転移学習を用いた方法を提案し、実験により評価する。キーワード 書誌情報抽出, ニューラルネットワーク, 参考文献文字列, word2vec, 能動サンプリング

1. はじめに

多数の学術論文を蓄積する電子図書館のサービスを利用する際、検索や文書間リンク等の機能は必須であり、これらの機能を利用するには、著者名やタイトルといった書誌情報が必要となる。しかし、これらの書誌情報を人手でデータベースに入力するコストは膨大なため、その作業を可能な限り自動で行う文書解析技術が求められている。そこで本稿では、学術論文の参考文献文字列に着目する。学術論文の参考文献欄には、多くの関連文献が記述されており、その著者やタイトルといった有用な書誌情報がある。浪越ら [1] は、Bi-directional LSTM-CNN-CRF を利用して、論文中の参考文献文字列から書誌情報を自動で抽出する方法を提案した。しかしこの手法で書誌情報を高精度に抽出するには、参考文献文字列の書式が異なる学術論文ごとに、数百件以上の参考文献文字列を学習用データとして用意する必要があり、その生成コストは無視できない。そこで本研究では、書誌情報抽出結果の確信度を利用し、学習データを選別することで必要な学習データを削減する能動サンプリングと、擬似学習データを自動生成するデータ拡張、また参考文献文字列の書式が異なる他雑誌のデータについて学習した書誌情報抽出器の推定結果を利用して、書誌情報抽出精度の向上を図る方法を提案する。

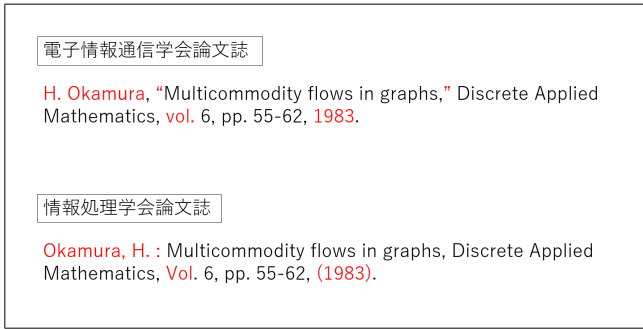
本稿の構成は次の通りである。まず、2. 節で学術論文からの書誌情報抽出に関する研究、また少量データによる機械学習の研究を紹介し、3. 節で本研究で用いる自動書誌情報抽出法について説明する。つづく 4. 節で少量学習データで書誌情報抽出精度を上げるための手法を提案し、実験方法とその結果を 5. 節で示し考察する。最後に 6. 節で本稿をまとめる。

2. 関連研究

2.1 書誌情報抽出に関する研究

多数の学術論文を格納する電子図書館では、書誌情報の管理は必須であり、学術論文から書誌情報を自動抽出する研究が行われている。ルールにより参考文献文字列から書誌情報を抽出する場合、例えば図 1 のように著者名、タイトル、発行年などの書式が異なる論文誌ごとに、書誌情報抽出のためのルールを設定する必要がある。電子図書館に格納される学術論文は増大しており、このようにルールを定義し、管理していくことは今後困難となることが予想される。

そこで、学習データを用意すれば利用可能な機械学習による書誌情報抽出が多く提案されている。例えば、CRF [2] を用いた書誌情報抽出に関する研究に、Peng ら [3], Council ら [4], Do ら [5], Cuong ら [6] の研究がある。Peng ら [3] は学術論文のタイトルページと参考文献欄から Title, Author 等それぞれ 13 項目の書誌情報を抽出する実験を行った。HMM, CRF, SVM により書誌情報抽出を行った結果、CRF が最も高い精度を実現しており、平均 F 値はタイトルページからの抽出では 0.939, 参考文献文字列からの抽出では 0.915 であった。一方、Council ら [4] は訓練済みの CRF モデルを利用して参考文献文字列から書誌情報を抽出する“ParsCit”を開発した。ParsCit は英文の参考文献文字列をトークン列に変換し、そのトークン列に対して書誌要素ラベルを付与する。実験では 3 つのデータセットにおいて参考文献文字列から書誌情報を抽出しており、Cora データセット [7] を対象にした実験では、平均 F 値は 0.950 であった。Do ら [5] は CRF により著者とその所属機関を識別する情報抽出システムである Enlil を開発した。ACL Anthology, ACM Digital Library, Cross Disciplinary Corpus を対象に著者名を抽出し、その所属機関



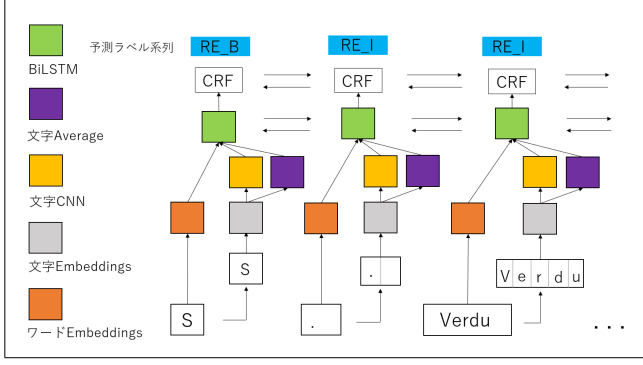


図3 浪越らのBi-directional LSTM-CNN-CRFモデル [1]

タに対応するラベルを24種類定義している [1]。

3.3 Bi-directional LSTM-CNN-CRF

本研究では参考文献書誌情報抽出に、Maら [10] の提案した Bidirectional LSTM-CNN-CRF を浪越らが独自に改良したモデルを用いる。モデルの構造を図3に示す。浪越らの Bi-directional LSTM-CNN-CRF は、まず word2vec [11,12] を用いて参考文献文字列をデリミタで分解して得たワードの分散表現を獲得し、ワードベクトルとワードの各文字のベクトルを得る。ここで、文字ベクトルは、keras [13] に実装されている Embedding レイヤーを用いて獲得する。次にワードの各文字のベクトルを CNN に入力として与え、固定次元の文字 CNN ベクトルを得る。さらにここで、ワードの各文字のベクトルの平均を求め、文字 Average ベクトルを算出する。次に、双方向の LSTM である BiLSTM にワードベクトルと文字 CNN ベクトル、文字 Average ベクトルを入力として与える。最後に BiLSTM 層の出力ベクトルを CRF に入力として与える。

4. 学習データの削減手法

4.1 能動サンプリング

4.1.1 能動サンプリングの概要

能動サンプリングでは、ある時点のモデルで、書誌情報抽出が困難な参考文献文字列を優先的に選択して学習データとし、逐次学習モデルを更新する。具体的にはまず、参考文献文字列 S を大量に収集する。次に、少量の参考文献文字列 $S_0 \subset S$ を選出して、これを第1回目の学習データとしてモデル M_0 を学習する。その後、以下の手順を繰り返す。モデル M_{t-1} を用いて、参考文献文字列 $S - \cup_{i=0}^{t-1} S_i$ の確信度をそれぞれ算出し、確信度の昇順にランキングする。その後、上位 n 件の参考文献文字列 $S_t \subset S - \cup_{i=0}^{t-1} S_i$ を選出する。そして $\cup_{i=0}^t S_i$ を学習データとしてモデル M_t を学習する。これは書誌情報抽出が困難なサンプルは学習に有効であるという考え方に基づく。

4.1.2 確信度

書誌情報抽出の困難さを表す尺度として、川上ら [14] の提案した確信度 Minimum Probability of Token Assignment (MP) を利用する。川上らの用いた確信度 MP が、各トークンに付与されたラベルの周辺確率を利用していたのに対して、本稿では、参考文献文字列中の各ワードに付与された書誌情報 BI ラベルの周辺確率を利用する。入力ワード系列 \mathbf{x} に対して、 Y_i を参

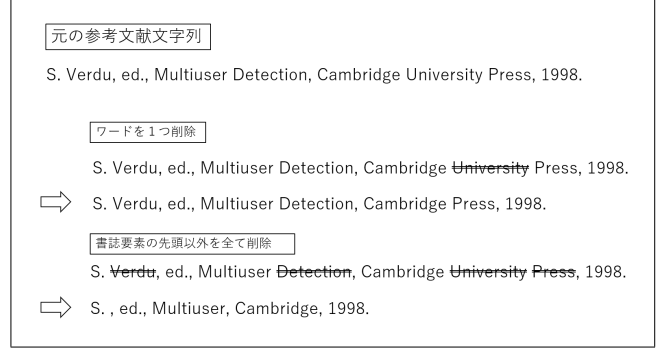


図4 擬似学習データの例

考文献文字列中の i 番目のワード x_i に対して付与されるラベルを表す確率変数とする。また L を付与できるラベルの集合とすると、式 (1) で表される確率は i 番目のワードに着目したラベル付与の確信度とみなすことができる。

$$\max_{l \in L} P(Y_i = l | \mathbf{x}) \quad (1)$$

そして、参考文献文字列中の各ワードに対するラベル付与の確信度の中で最小のものを、その参考文献文字列の書誌情報抽出の確信度とする。具体的には以下の式で定義される。

$$C_{MP}(\mathbf{x}) = \min_{1 \leq i \leq |\mathbf{x}|} \max_{l \in L} P(Y_i = l | \mathbf{x}) \quad (2)$$

4.2 データ拡張

データ拡張により生成する擬似学習データを利用して抽出精度の向上を図る方法について説明する。本研究においてデータ拡張は、参考文献文字列中のワードを無作為に n 件削除することによって行う。本手法では、参考文献文字列中の書誌要素の先頭に該当するワードは削除せず、書誌要素の先頭以外に該当するワードを無作為に選出し、削除する。生成された擬似学習データの例を図4に示す。

4.3 転移学習

通常、雑誌ごとに参考文献文字列の書式は異なるため、高精度に書誌情報を抽出するためには対象雑誌の学習データを用いて書誌情報抽出器を学習する。よって他雑誌のラベル付きデータが大量にあったとしてもそれらは利用されない。しかし、参考文献文字列に現れる特徴、また抽出する書誌情報には書式の異なる雑誌においても共通点が存在する。そこで、これらの情報を対象雑誌の書誌情報抽出器で利用することができれば、書誌情報抽出精度の向上が期待できる。本研究では他雑誌の書誌情報抽出器を用いて対象雑誌の書誌情報抽出精度を向上させる方法を検討し、実験により検証する。具体的には、対象雑誌のラベル付き学習データが少量の場合に、他雑誌で学習された書誌情報抽出器のパラメータを初期値として、抽出対象雑誌の学習データでファインチューニングし、抽出対象雑誌の書誌情報を抽出する。

5. 評価実験

5.1 実験環境

少量学習データによる書誌情報の抽出精度を検証するため、

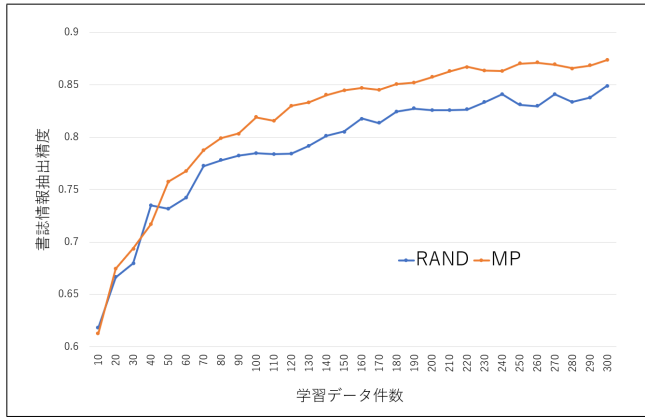


図 5 能動サンプリングによる書誌情報抽出結果

評価実験を行う。実験データとして、2000 年の電子情報通信学会英文論文誌 (IEICE-E) に含まれる参考文献文字列 4,497 件を利用する。また、書誌要素として、表 1 に示した 18 種類のラベルを定義したが、評価の際には浪越らに倣い、RA, RE, RTR, RAOT を AUTHOR, RT, RBT を TITLE, RW, RC を JOURNAL, RV, RN, RPP を VOLUME, RP を PUBLISHER, RD を DAY, RM を MONTH, RY を YEAR, RL, RURL, ROT を OTHER と再分類したものをを用いる [1]。つまり再分類したラベルが同じものは正解判定において区別しない。本研究では参考文献文字列の書誌要素を構成する全てのワードに正しい書誌要素ラベルを付与した場合、その参考文献文字列の書誌要素推定に成功したと判定する。なお、デリミタを構成するワードの正解判定は行わない。実験結果に示す書誌情報抽出精度は、推定に成功した参考文献文字列の割合を表す。また、書誌情報抽出精度は 5 分割交差検定で算出する。

5.2 能動サンプリングの有効性評価

まず、4.1 節で説明した能動サンプリングにより、書誌情報抽出精度を維持したまま、必要な学習データ量がどの程度削減できるかを実験により検証する。5 分割交差検定を行うため、雑誌中の参考文献文字列を 5 つに分割し、そのうち 4 つを学習データ、残りの 1 つをテストデータとする。実験では能動サンプリングにおける初期学習データ数と、追加する学習データ件数はともに 10 件とする。また、比較のため、学習データとして参考文献文字列を無作為に 10 件ずつ選出した場合を RAND と記し、ベースラインとする。

図 5 に能動サンプリングと RAND の結果を示す。図 5 ではデータ件数が 40 件を超えてからは常に MP が RAND の書誌情報抽出精度を上回っており、例えば学習データ件数が 180 件の時点で書誌情報抽出精度は 85% に到達した。RAND の書誌情報抽出精度が 85% に到達したのは学習データ件数が 300 件の時点であり、MP は RAND に比べて 120 件少ない学習データ件数で、同じ書誌情報抽出精度を得ていることから、確信度 MP による能動サンプリングは、Bi-directional LSTM-CNN-CRF による書誌情報抽出において有効であるといえる。

5.3 データ拡張の有効性評価

本節では 4.2 節で説明したデータ拡張により、書誌情報抽出

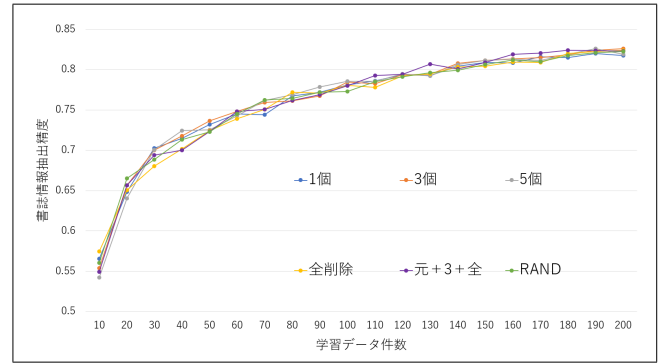


図 6 データ拡張による書誌情報抽出結果

精度がどの程度向上するかを実験により検証する。実験では学習データを各学習データ件数ごとに全て無作為に選出する。RAND は学習データを無作為に選出し、擬似学習データを加えない場合であり、これをベースラインとする。また、ベースライン、提案手法ともに 5 分割交差検定を 3 回行って結果の平均を算出し、その値を書誌情報抽出精度とする。学習はデータ拡張を行う前の元の学習データにデータ拡張により生成された擬似学習データを追加して行う。データ拡張で得られる擬似学習データは元の参考文献文字列中のワードを 1 個、3 個、5 個、また書誌要素の先頭以外のワードを全て削除 (全削除) した文字列である。つまり本節の実験における学習データは「元の参考文献文字列 + ワードを k 個削除した参考文献文字列」となる。これらに加えて、「元の参考文献文字列 + 3 個削除した参考文献文字列 + 全削除した参考文献文字列」を学習データとした場合についても実験を行う。

実験結果を図 6 に示す。ワードを 5 個削除した擬似学習データを加えた場合、学習データを「元の参考文献文字列 + 3 個削除した参考文献文字列 + 全削除した参考文献文字列」とした場合については、学習データ件数が 10 件から 200 件の間の全 20 回の RAND と比較して、計 14 回書誌情報抽出精度が高くなった。さらに、学習データを「元の参考文献文字列 + 3 個削除した参考文献文字列 + 全削除した参考文献文字列」とした場合は、学習データ件数が 100 件を超えてからは常に RAND よりも高い精度で推移しており、書誌情報抽出精度が 0.8 に達したのは学習データ件数が 130 件の時で、図 6 の中で最も早かった。また、RAND と比較して、ワードを 1 個削除した擬似学習データを加えた場合は 12 回、3 個削除した擬似学習データを加えた場合は 13 回書誌情報抽出精度が高くなっている。それに対して削除できるワードを全て削除した擬似学習データを加えた場合は書誌情報抽出精度が RAND よりも高くなったのは 8 回という結果になった。

5.4 転移学習の有効性評価

5.4.1 他雑誌用書誌情報抽出器による書誌情報抽出精度

本実験では、5.1 節で説明した参考文献文字列コーパスに加えて、転移学習のために 1952 年から 2012 年までの IEEE Trans. Computers (IEEE-CS) に含まれる参考文献文字列の引用回数上位 4,770 件を利用する。まず、他雑誌用書誌情報抽出器による書誌情報抽出精度を実験により検証する。実験は IEICE-E,

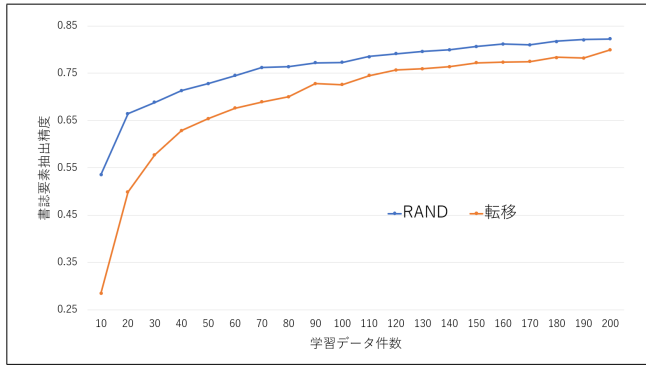


図 7 転移学習の結果 (抽出対象 IEICE-E, 転移雑誌 IEEE-CS)

IEEE-CS の 2 雑誌を対象にそれぞれ 4000 件弱の学習データを用いて各書誌情報抽出器を学習し、書誌情報を抽出する。その書誌情報抽出精度を表 2 に示す。抽出対象雑誌と異なる雑誌で学習を行った場合、抽出対象と同じ雑誌で学習を行った場合と比べて書誌情報抽出精度が低くなっていることが分かる。特に抽出対象が IEEE-CS、学習データが IEICE-E のとき、書誌情報抽出精度は 0.200 と著しく低くなった。このことから書誌情報を高精度に抽出するには、抽出対象雑誌と同じ雑誌で学習を行うことが必須であると分かる。

5.4.2 転移学習の効果

本節では 4.3 節で説明した転移学習により、書誌情報抽出精度がどの程度向上するかを実験により検証する。実験では 5.3 節と同様に、学習データを各学習データ件数ごとに全て無作為に選出する。ベースラインとする無作為抽出である RAND、転移学習ともに 5 分割交差検定を 3 回行って結果の平均を算出し、書誌情報抽出精度とする。なお、他雑誌の書誌情報抽出器は、それぞれの雑誌の全参考文献文字列の五分の四 (4,000 件弱) を学習データとして学習したモデルである。

IEEE-CS の 4,000 件弱の参考文献文字列で学習した書誌情報抽出器を、抽出対象雑誌である IEICE-E で無作為に選んだ 10 件から 200 件の参考文献文字列でファインチューニングをして書誌情報抽出を行った結果を図 7 に示す。図 7 より転移学習を行った場合、RAND よりも全体的に書誌情報抽出精度が低くなった。特に、学習データ件数が 80 件以下のとき、その差は大きくなっていることがわかる。ここで、学習データ件数 50 件の時点での各書誌要素の抽出精度を表 3 に示す。表 3 より書誌要素“YEAR”については RAND と比較して抽出精度が高くなっていることがわかる。それに対し、ほかの書誌要素の抽出精度は RAND とほぼ同等か、それ以下という結果だった。

また、RAND と転移学習の書誌情報抽出精度の差が最も大きくなった学習データ件数 10 件の時点での各書誌要素の抽出

表 3 学習データ 50 件時点での各書誌要素の抽出精度

	RAND	転移
AUTHOR	0.9797	0.9784
TITLE	0.9795	0.9672
JOURNAL	0.9472	0.9455
VOLUME	0.9506	0.9490
PUBLISHER	0.6401	0.6240
DAY	0.4074	0.3647
MONTH	0.9455	0.9213
YEAR	0.9476	0.9505
OTHER	0.4873	0.4836

表 4 学習データ 10 件時点での各書誌要素の抽出精度

	RAND	転移
AUTHOR	0.9767	0.9709
TITLE	0.9465	0.9231
JOURNAL	0.9201	0.7842
VOLUME	0.9374	0.8415
PUBLISHER	0.3318	0.3363
DAY	0.1285	0.0350
MONTH	0.7972	0.5846
YEAR	0.9389	0.8265
OTHER	0.1768	0.2313

精度を表 4 に示す。表 4 より RAND、転移学習ともに書誌要素“DAY”の抽出精度が他の書誌要素に比べて極端に悪いことがわかる。しかし、書誌要素“DAY”は IEICE-E の参考文献文字列 4,497 件の中に 109 件しか出現しない書誌要素であるため、他の書誌要素に比べて、参考文献文字列全体の抽出精度に与える影響は小さい。また、書誌要素“AUTHOR”と“TITLE”については RAND とほぼ同等の抽出精度となっており、“PUBLISHER”、“OTHER”に関しては RAND の抽出精度を上回った。RAND と比較して最も抽出精度が悪かった書誌要素“MONTH”には、IEICE-E では「Sept.」のように記述されるのに対して、IEEE-CS では「September」のように記述されるという書式の違いがある。この違いが書誌要素“MONTH”の抽出精度を下げた要因としてあげられる。

さらに、RAND と転移学習の書誌情報抽出精度の差が最も小さくなった学習データ件数 200 件の時点での各書誌要素の抽出精度を表 5 に示す。表 5 より“PUBLISHER”の抽出精度は転移学習の方が RAND よりも高くなり、その他の書誌要素の抽出精度についても RAND には及ばないものの、ほぼ同等という結果になった。ここで、RAND と比較して最も抽出精度が悪かった書誌要素“OTHER”について考察する。“OTHER”は参考文献文字列中のロケーション、URL などの書誌要素にあたる。IEICE-E、IEEE-CS の 2 雑誌の“OTHER”の書式の違いについて図 8 に示す。図 8 中の赤字で書かれたトークンが参考文献文字列中のロケーション、つまり“OTHER”にあたるトークンである。この例において、IEICE-E ではロケーションにあたる書誌要素が“YEAR”の直前にあるのに対して、IEEE-CS では“YEAR”の後の学会名に続く位置にある。また、IEICE-E の参考文献文字列ではこのロケーションのほ

表 2 他雑誌用抽出器による書誌情報抽出精度

抽出対象データ	書誌情報抽出器の学習データ	精度
IEICE-E	IEICE-E	0.922
	IEEE-CS	0.836
IEEE-CS	IEICE-E	0.200
	IEEE-CS	0.862

表 5 学習データ 200 件時点での各書誌要素の抽出精度

	RAND	転移
AUTHOR	0.9903	0.9833
TITLE	0.9829	0.9797
JOURNAL	0.9684	0.9663
VOLUME	0.9728	0.9702
PUBLISHER	0.7582	0.7865
DAY	0.7798	0.7638
MONTH	0.9674	0.9552
YEAR	0.9681	0.9675
OTHER	0.7333	0.7156

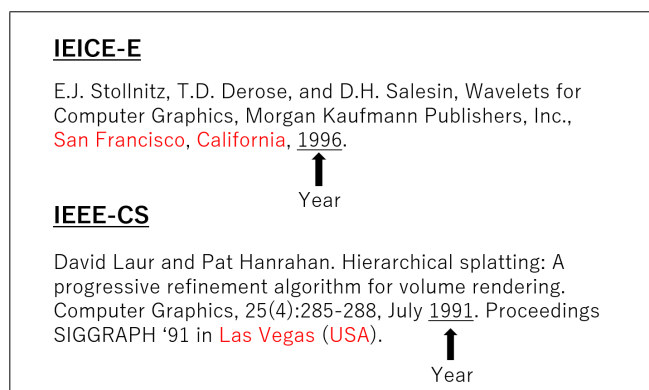


図 8 IEICE-E と IEEE-CS による “OTHER” の書式の違い

ば全てが図 8 の例のような位置に書かれる傾向があるのに対し、IEEE-CS に含まれる参考文献文字列では、ロケーションが“YEAR”の直前や直後、また直前よりもさらに前に書かれるなど様々であった。これが、特にこの“OTHER”について、転移学習が有効でなかった理由の一つにあげられる。

6. ま と め

本稿では、Bi-directional LSTM-CNN-CRF による参考文献書誌情報抽出において、能動サンプリング、データ拡張、転移学習を利用して、学習データ件数と書誌情報抽出精度の関係を実験により検証した。その結果まず、参考文献文字列をデリミタで分割して得た各ワードに付与された書誌情報 BI ラベルの周辺確率に基づく確信度 MP を利用した能動サンプリングでは、無作為に学習データを増やした場合の学習データ件数 300 件での書誌情報抽出精度に、学習データ件数 180 件で到達した。参考文献文字列中のワードを削除して擬似学習データとするデータ拡張では、ワードを 5 個削除した擬似学習データを追加した場合と、学習データを「拡張前の参考文献文字列 + ワードを 3 個削除して生成した擬似学習データ + 書誌要素の先頭以外を全て削除して生成した擬似学習データ」とした場合には、他のワード数で擬似学習データを生成し学習データに追加した場合に比べ、書誌情報抽出精度が向上した。また、抽出対象雑誌と異なる雑誌で学習を行った他雑誌用書誌情報抽出器を、抽出対象雑誌の学習データ 10 件から 200 件でファインチューニングし、抽出対象雑誌の書誌情報抽出を行った転移学習においては書誌情報抽出精度は向上しなかった。

今後の課題としては、参考文献文字列中のワードやトーク

ンを逆翻訳することにより擬似学習データを生成する Back-Translation によるデータ拡張の実装、転移学習の方法の改善などがあげられる。

謝 辞

本研究の一部は、科学研究費補助金基盤研究 (C)(課題番号 18K11989)、および新エネルギー・産業技術総合開発機構 (NEDO) の戦略的イノベーション創造プログラム (SIP) 第二期「ビッグデータ・AI を活用したサイバー空間基盤技術」および国立情報学研究所共同研究の援助による、ここに記して感謝する。

文 献

- [1] 浪越大貴, 太田学, 高須淳宏, 安達淳, Bi-directional LSTM-CNN-CRF による参考文献書誌情報抽出,” 信学技法, vol. 118, no. 377, pp. 17-22, 2018.
- [2] J. Lafferty, A. McCallum, and F. Pereira, “Conditional Random Fields : Probabilistic Models for Segmenting and Labeling Sequence Data,” in Proc. of ICML 2001, 2001, pp. 282-289.
- [3] F. Peng, and A. McCallum, “Accurate Information Extraction from Research Papers Using Conditional Random Fields,” in Proc. of HLT-NAACL 2004, pp. 329-336, 2014.
- [4] I. G. Councill, C. L. Giles, and M. Y. Kan, “ParsCit: an open-source CRF reference string parsing package,” in Proc. of LREC 2008, pp. 661-667, 2008.
- [5] H. H. N. Do, M. K. Chandrasekaran, P. S. Cho, and M. Y. Kan, “Extracting and matching authors and affiliations in scholarly documents,” in Proc. of JCDL 2013, pp. 219-228, 2013.
- [6] N. V. Cuong, M. K. Chandrasekaran, M. Y. Kan, and W. S. Lee, “Scholarly Document Information Extraction using Extensible Features for Efficient Higher Order Semi-CRFs,” in Proc. of JCDL 2015, pp. 61-64, 2015.
- [7] A. McCallum, K. Nigam, J. Rennie, and K. Seymore, “Automating the Construction of Internet Portals with Machine Learning,” Information Retrieval, vol. 3, no. 2, pp. 127-163, 2000.
- [8] M. Ohta, R. Inoue, and A. Takasu, “Empirical Evaluation of Active Sampling for CRF-Based Analysis of Pages,” in Proc. of IEEE IRI 2010, pp. 13-18, 2010.
- [9] W. Y. Wang ,and D. Yang, “That’s so annoying!!!: A lexical and frame-semantic embedding based data augmentation approach to automatic categorization of annoying behaviors using #petpeeve tweets,” in Proc. of Conference on Empirical Methods in Natural Language Processing 2015, pp. 2557-2563, 2015.
- [10] X. Ma, and E. Hovy, “End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF,” in Proc. of 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016) , pp. 1064-1074, Association for Computational Linguistics, 2016.
- [11] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in Proc. of 26th International Conference on Neural Information Processing System (NIPS 2013), vol. 2, pp. 3111-3119, 2013.
- [12] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” CoRR, vol.abs/1301.3781, pp.1-12, 2013. <http://arxiv.org/abs/1301.3781>
- [13] F. Chollet, “Keras,” 2015. <http://keras.io>
- [14] 川上尚慶, 太田学, 高須淳宏, 安達淳, “少量学習データによる

参考文献誌情報抽出精度の向上,” 情報処理学会論文誌データベース, vol. 8, no. 2, pp. 18-29, 2015.