Method to Predict Confidential Words in Japanese Judicial Precedents Using Neural Networks With Part-of-Speech Tags

Masakazu Kanazawa, Atsushi Ito, Kazuyuki Yamasawa, Takehiko Kasahara, Yuya Kiryu and Fubito Toyama

***Abstract*—Cognitive Infocommunications involve a combination of informatics and telecommunications. In the future, infocommunication is expected to become more intelligent and life supportive. Privacy is one of the most critical concerns in infocommunications. Encryption is a well-recognized technology that ensures privacy; however, it is not easy to completely hide personal information. One technique to protect privacy is by finding confidential words in a file or a website and changing them into meaningless words. In this paper, we investigate a technology used to hide confidential words taken from judicial precedents. In the Japanese judicial field, details of most precedents are not made available to the public on the Japanese court web pages to protect the persons involved. To ensure privacy, confidential words, such as personal names, are replaced by other meaningless words. This operation takes time and effort because it is done manually. Therefore, it is desirable to automatically predict confidential words. We proposed a method for predicting confidential words in Japanese judicial precedents by using part-of-speech (POS) tagging with neural networks. As a result, we obtained 88% accuracy improvement over a previous model. In this paper, we describe the mechanism of our proposed model and the prediction results using perplexity. Then, we evaluated how our proposed model was useful for the actual precedents by using recall and precision. As a result, our proposed model could detect confidential words in certain Japanese precedents.**

***Index Terms*—confidential word, neural network, Part of Speech (POS) tag, perplexity (PPL), precision, recall**

I. INTRODUCTION

A. Cognitive Infocommunications

Cognitive Infocommunications (CogInfoCom) [1][2] involves a combination of informatics and communications. CogInfoCom systems extend human cognitive capabilities by providing fast infocommunications links to huge repositories of information produced by the shared cognitive activities of social communities [3]. CogInfoCom is expected to become more intelligent, and it would even have the ability to support life. Fig. 1 shows the idea of CogInfoCom. Clearly, privacy is one of the most critical concerns in infocommunications. Encryption is a well-recognized technology used for ensuring privacy; however, encryption does not effectively hide personal information completely. One technique to protect privacy is to determine the confidential words in a file or a website and convert them into meaningless words. CogInfoCom makes a network intelligent and automatically changes confidential words into meaningless words.

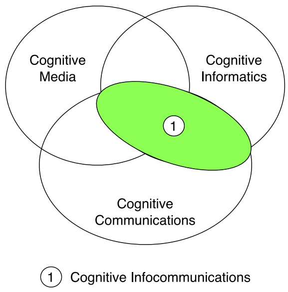

Fig. 1. Infocommunication model

B. IT-Based Court: Cyber Court

Globalization of the economy, international trade, and disputes present new demands on judiciaries worldwide. At the same time, advances in information communication technology (ICT) offer opportunities to judicial policymakers to make justice more accessible, transparent, and effective.

By introducing ICT, many countries have allowed easy access judicial documents easily. Such a justice system empowered by ICT is called a "cyber court." A pioneering study of a cyber court system is Courtroom 21 [4], which started in 1993 in the College of William & Mary as a joint project between the university and the National Center for State Courts in the United States of America.

In Japan, the prototype for the first civil trial was developed in the Toin University of Yokohama in 2004 [5, 6], and its effectiveness was proved particularly to the Japanese citizen judge system [7]. An experiment with a remote trial was also conducted [8]. The Investments for the Future Strategy 2017 by the Japanese Cabinet Office includes ICT conversion for trials to accelerate the trials and improve the efficiency of the judicial system [9].

Masakazu Kanazawa is with the Graduate School of Engineering Utsunomiya University, Tochigi, Japan (e-mail: kanama2591@gmail.com).
Atsushi Ito is Graduate School of Engineering Utsunomiya University, Tochigi, Japan (e-mail: at.ito@is.utsunomiya-u.ac.jp).
Kazuyuki Yamasawa is with the TKC Corporation, Tokyo, Japan (e-mail: yamasawa-kazuyuki@tkc.co.jp).
Takehiko Kasahara is the Toin Yokohama University, Kanagawa, Japan (e-mail: kasahara@toin.ac.jp).
Yuya Kiryu is with the KDDI Corporation, Tokyo, Japan (e-mail: u81645@gmail.com).
Fubito Toyama is with the Graduate School of Engineering Ursunomiya University,Tochigi,Japan (email: fubito@is.utsunomiya-u.ac.jp)

DOI: 10.36244/ICJ.2020.1.3

C. Predicting Confidential Words

To protect personal information, most precedents are not open to the public on the Japanese court web pages. Confidential words (e.g., personal, corporate, and place names) in open precedents are replaced by other meaningless words, such as a single uppercase letter "A." This operation takes time and effort because it is done manually. Therefore, we would like to predict confidential words automatically to solve this problem.

In recent years, the use of neural networks has advanced in natural language processing. The research includes deriving a vector by considering the word meanings and predicting words that are actively ongoing [10]. We reported earlier that a bi-directional long short-term memory (LSTM) with left–right (LR) (hereinafter Bi-directional LSTM–LR) model is effective to predict target words in Japanese precedents. However, we did not obtain good accuracy for the detection of confidential words [11]. In this research, we attempt to improve the accuracy of predicting confidential words. In the Japanese precedents, we found that the confidential words were mostly proper nouns and various parts of speech (POS). Therefore, we considered a new method by using a POS tag.

In this paper, we propose a new method using a neural network combined with the POS tag to improve accuracy, and we describe the experimental results for predicting the confidential words. Then, we show the probability of applying our model to practical situations.

II. HOW TO ANONYMIZE CONFIDENTIAL WORDS

Japanese precedents include many confidential words, such as personal names, corporate names, and place names. To protect privacy, such words are converted into meaningless words. In Japan, this procedure takes time and effort because it is done manually.

A. Problem of Anonymizing Confidential Words

Some judicial precedents are available on the website of the court [12]. Confidential words in these precedents have been replaced with a letter of the alphabet. (In paid magazines and websites, Japanese letters are sometimes used.) Fig. 2 shows an example of a replaced word. This process is performed manually. These replacements cannot be done easily by using a dictionary of proper nouns because confidential words sometimes have multiple meanings, and it is difficult to distinguish among them. Therefore, the substitutions are done manually by legal experts based on the context.

B. Aim of Our Study

Our purpose is to extract the confidential words in Japanese precedents and automatically replace the confidential words with a single letter of the English alphabet ("A"). There are various methods for predicting confidential words in the judicial precedents. One possibility is to use a dictionary of proper nouns. However, even if the confidential word matched the list in the dictionary, the word may sometimes be used with a different meaning in the Japanese precedent. For example, the word "Yamaguchi" may refer to a city or a person. To solve this problem, we will propose a method for predicting confidential words in a sentence based on the context by using a neural network. Fig. 3 shows the prediction mechanism of the confidential words.

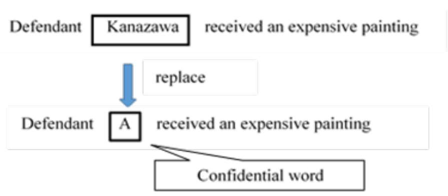

Fig. 2. Anonymizing the confidential word

For the preprocessing, we converted the confidential words contained in the datasets to the uppercase letter "A" and separated the Japanese words with spaces by using MeCab, a Japanese morphological analyzer [13]. When the Japanese precedents (corpus) containing the confidential words replaced by "A" are entered into the neural network, they are learned by the neural network, which then predicts the confidential words.

C. Related Works

Named-entity extraction is a widely used technique to obtain the target words in a sentence. Named-entity recognition (NER) is probably the first step for information extraction to locate and classify the named entities in the text into predefined categories, such as the names of persons, organizations, locations, expressions of times, quantities, monetary values, and percentages. NER is used in many fields in natural language processing [14] [15].

NER extraction is executed mostly by using two methods: the rule-based method (by pattern matching) and the statistical method (by machine learning). The method using pattern matching has a very high cost because the pattern of the named entity dictionary needs to be created and updated manually. Various machine learning methods have been studied to solve the problem. Machine learning methods, such as the hidden Markov model and conditional random fields (CRFs), can learn the pattern of a named entity by preparing the corpus. CRF has proved to be quite successful for NER. Nevertheless, the problem of machine learning is that the cost of manually making a corpus is quite high [16].

III. RESULTS OF PREVIOUS STUDIES

The use of neural networks is widespread even for learning natural languages. Therefore, we will study a neural network to predict the confidential word because the embedding vector of the word used in the neural network is very effective to

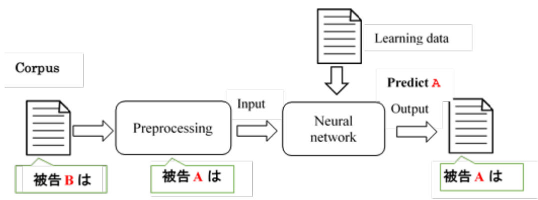

Fig. 3. Prediction mechanism for confidential words

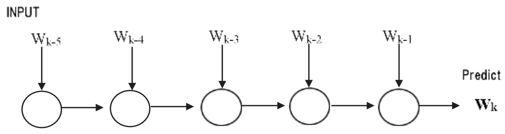

Fig. 4. LSTM model

handle word meanings. We investigated some models of the neural network as described in the following subsections.

A. LSTM Model

A neural network is beneficial in the field of natural language learning. From the words adjacent to the target word, we can decide whether or not the target word is confidential. We found that the most effective model was LSTM), which was an improved model of the recurrent neural network (RNN). The RNN suffers from vanishing or exploding gradient or the exploding problem when the input data is long. LSTM is very useful in long sequential data (Fig. 4) [17].

B. Bi-Directional LSTM-LR Model

We used the Bi-directional LSTM–LR model to imitate the anonymization work done by humans. When humans do this work, they always judge the word after reading the words on the left and right of the target words. We proposed this model at first, as shown in Fig. 5.

For the input order on the backward (right side) of the target word in the reverse order of the sentence, we assume that the influence of the target word increases with increasing proximity of the input word to the target word.

C. Corpus and Experiment of the Bi-Directional LSTM–LR Model

We used 50,000 judicial precedents for the training data and 10,000 judicial precedents for the test data. In these data, the contents of the trials held from 1993 to 2017 were recorded. These were the precedents database provided by the TKC Co. [18]. The various parameters used are shown in Table 1.

Window size means the chunk size, which describes the input word size before or after the target word. Fig. 5 shows the window size as 4 to explain the model; however, in this experiment, we used a window size of 10.

For accuracy, we used the perplexity (hereinafter PPL) that was used in previous studies for predicting the next word. In

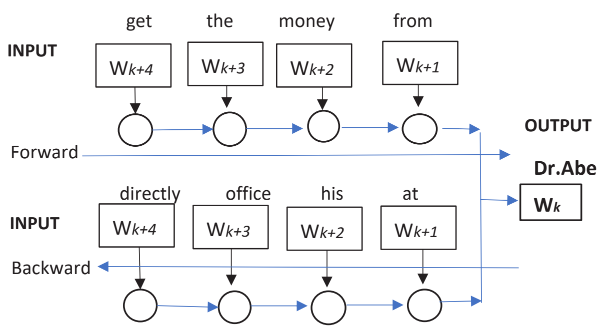

Fig. 5. Bi-directional LSTM–LR model

TABLE 1. PARAMETERS OF THE MODEL

Hidden layer	100
Embedding size	200
Window size	10
Batch size	200
Learning rate	0.001
Loss	Softmax function

TABLE II. Result of the experiment

	Simple LSTM	Bi-directional LSTM–LR
PPL	4.8	4.7
CW_PPL	56.3	37.3

natural language processing, PPL is usually used for evaluating the language model. PPL is defined as follows:

$$PPL = 2^{-\frac{1}{N}\sum_{i=1}^{N} log_2 p(w_i)} \quad (1)$$

In Eq. (1), p is the probability, and N is the total number of the words. PPL represents the number of prediction choices that are narrowed down to neural networks. The smaller the value, the better the prediction results.

D. Analysis of the Experiment Result

CW_PPL is the average PPL for the test data, which is the PPL of the confidential words, whereas PPL is the average for all the test data. The PPL scores are shown in Table Ⅱ.

Our experiment proved that our proposed model using neural networks was effective for predicting the target words. Nevertheless, the CW_PPL score was poor; therefore, the accuracy of predicting the confidential words needed to be improved. One reason was the possibility of paraphrasing words such as "plaintiff," "defendant," "doctor," and "teacher." These paraphrased words could be excluded from the choices based on the results of the calculated probability. We need to review the algorithm for preprocessing the input corpus before inputting the algorithm to the neural network because the difference in the scores between PPL and CW_PPL was too large.

IV. NEW MODEL TO IMPROVE ACCURACY

Our algorithm did not give good accuracy for predicting the confidential word; therefore, we investigated the algorithm of our model and reviewed it. In general, words have meanings and (POS) tags in the dictionary. We found almost all the confidential words had the POS tag of proper nouns. Therefore, if the POS tag is added to the neural network with the words, it may be possible to learn better and improve the accuracy. The most popular tagging tool for Japanese sentences is MeCab.

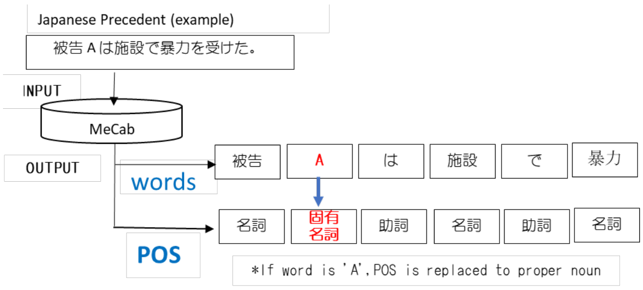

Fig. 6. POS tagging using MeCab

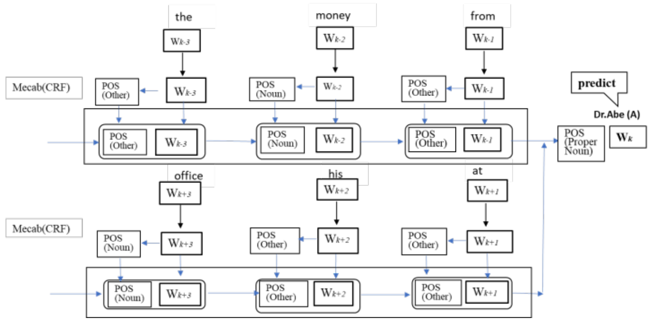

Fig. 7. Proposed model that combined the POS tag

A. MeCab as CRF

MeCab is the most powerful tool to extract the POS tag from the words in Japanese judicial precedents. It is a well-known Japanese morphological analyzer and is a kind of CRF. CRF is a successful named-entity extraction output technique used to label information, such as POS tagging. Japanese sentences have no spaces between words; therefore, MeCab inserts spaces between words and tags the words (using POS). If the word is "A," it is a confidential word. The POS corresponding to "A" is replaced by the proper noun described in Fig. 6.

B. Bi-Directional LSTM–LR Combined With the POS Tag

In our previous model, the corpus (words) were input into the neural network as done for natural language processing. To improve the CW_PPL score, we attempted to input the POS tag corresponding to the word extracted by MeCab (CRF) in the previous model (Bi-directional LSTM–LR). The outline of this proposed model is shown in Fig. 7, and the identification mechanism is shown in Fig. 8. Summary of the algorithm is described as bellows and detail is given in the appendix.

Input data

$$\boldsymbol{L} = (\mathrm{w_i})_i^{bj} \quad : w_1, w_2, \cdots w_{10} \quad \text{:word(backward)} \qquad (2)$$

$$\boldsymbol{L'} = (\mathrm{w_i})_i^{fj} : w_{-1}, w_{-2}, \cdots w_{-10} \quad \text{:word(forward)} \qquad (3)$$

$$\boldsymbol{P} = (p)_i^{bj} \quad : p_1, p_2, \cdots p_{10} \quad \text{:POS(backward)} \qquad (4)$$

$$\boldsymbol{P'} = (p_i)_i^{fj} \quad : p_{-1}, p_{-2}, \cdots p_{-10} \quad \text{:POS(forward)} \qquad (5)$$

$$\text{Output } \boldsymbol{O_i} = \text{LSTM } (\boldsymbol{L} + \boldsymbol{L'} + \boldsymbol{P} + \boldsymbol{P'}) \quad \textbf{\textit{:output}} \qquad (6)$$

MeCab is different from conventional natural language processing technology. Each W*iki* word describes a word, and every other word or every noun is the associated POS tag of the W*ki* word. When the Japanese precedents (corpus) is input into MeCab, MeCab separates the words by inserting spaces and tags them as a POS. If the confidential word ("A") appears in the precedents, then "proper noun" is tagged to the confidential word. Next, we assign a unique index to the POS tag and make a part dictionary by merging them into the input data (word) for the new neural network using an embedding vector.

C. Experiment of the Proposed Model Combined With the POS Tag

We experimented using the proposed model combined with the CRF. We used 10,000 judicial precedents for the training data and 5,000 judicial precedents for the test data from 2013 to 2017. The data was the same as the previous one. However, the number of training data sets was smaller because the data became approximately two times larger than the previous data by adding the POS tag. The various parameters were the same as that in Table Ⅰ. We used the same evaluation method as that in our previous experiment.

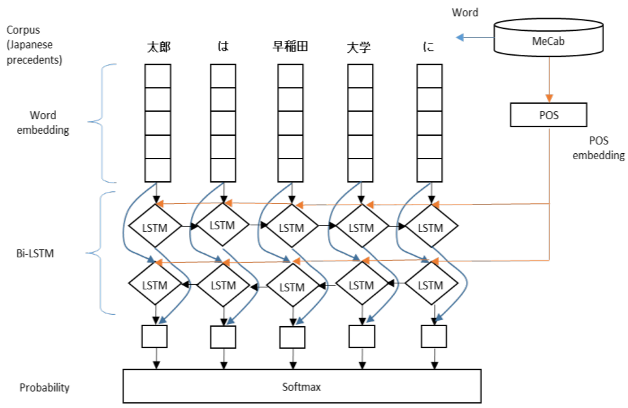

Fig.8. Identification mechanism

The results of this experiment as compared with the results of the Bi-directional LSTM–LR mode using the same input corpus are shown in Table Ⅲ.

The PPL score showed a 23% improvement in accuracy over the previous model (Bi-directional LSTM–LR), and the CW_PPL score also showed a 30% improvement in accuracy.

Therefore, we found that the Bi-directional LSTM–LR model, which combined the words and POS tag, was very effective in predicting the confidential words. However, the CW_PPL score needed further improvement.

D. Improving the Preprocessing Algorithm

Before learning the input corpus by using neural networks, it is necessary to reprocess the corpus.

TABLE Ⅲ. RESULT OF THE SECOND EXPERIMENT

	Proposed model	Bi-directional LSTM-LR
PPL	4.1	5.2
CW_PPL	28.6	40.7

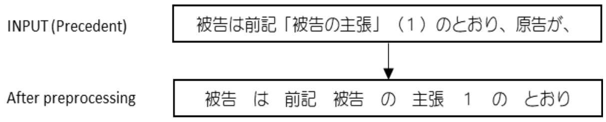

Fig.9. Example of the preprocessing of Japanese precedents

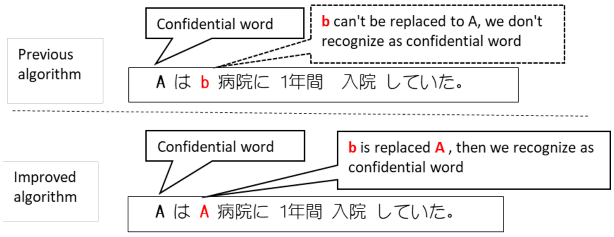

Fig. 10. Example of the improved algorithm

For example, many punctuation marks used for separating words and phrases, such as「」and（）, appear often in Japanese judicial precedents. These marks are only noise for learning; therefore, we omitted them (Fig. 9).

However, punctuation mark "。" was not omitted to prevent the flow of the context. The preprocessing was also done in the first experiment.

In Japanese precedents, confidential words are replaced by not only half-width uppercase letters but also by full-width uppercase and lowercase letters. In the previous algorithm, only when a single half-width capital letter of the alphabet appeared in the precedent, we replaced it with the half-width uppercase letter "A." Therefore, when the lowercase letter "b" appeared in the precedent, it could not replace "A"; therefore, we did not recognize "b" as a confidential word (see Fig. 10).

Therefore, we improved this algorithm by stating that if both single half-width and full-width letters appeared in the precedent, we replaced it by a half-width uppercase capital letter "A" as the confidential word.

E. Experiment Result After Using Improved Algorithm

The results of the experiment after using the improved algorithm are shown in Table Ⅳ.

TABLE Ⅳ. RESULT OF THE EXPERIMENT AFTER USING THE IMPROVED ALGORITHM

	New proposed model after using the improved algorithm	Previous model (Bi-directional LSTM–LR)
PPL	4.1	5.2
CW_PPL	5.1	40.7

CW_PPL showed that after using the improved algorithm, the proposed model accuracy decreased by 35.6 as compared with the previous model (Bi-directional model). We found that the CW_PPL score had improved 88% in accuracy as compared with the previous model.

As a result, we confirmed that our proposed model (Bi-directional LSTM–LR combined CRF) had high accuracy for predicting confidential words.

We got excellent predicting ability with our proposed model; therefore, we needed to confirm whether the model would be practical or not; this is our next step.

V. EVALUATION OF THE PROPOSED MODEL FOR APPLICATION

In an actual legal record, it is essential to evaluate whether the confidential word can be correctly recognized or whether the non-confidential word has also been recognized as a confidential word.

A. Training the Proposed Model Using Anonymized Precedents

We evaluated how our proposed model affects some types of anonymized precedents. We used the following parameters to examine the accuracy.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

TABLE Ⅴ. RESULTS OF THE EXPERIMENT IN VARIOUS TYPES OF ANONYMIZED PRECEDENTS

Item	Total words	Confidential word			Normal word			**Recall**	**Precision**	**F1**
		Appear (TP+FN)	No-hit (FN)	Hit (TP)	Appear (TN)	Hit (FP)	No-hit			
Rental contact	7800	52	29	23	7748	623	7125	44%	4%	7%
Land contract	9600	1588	751	837	8012	806	7206	53%	51%	52%
Traffic accident	2600	67	10	57	2533	280	2253	85%	17%	28%
Traffic accident	8400	100	35	65	8300	831	7469	65%	7%	13%
Rental contract	4000	76	11	65	3924	250	3674	86%	21%	33%
Injury case	12800	543	169	374	12257	1624	10633	**69%**	19%	29%
Land contract	1800	34	26	8	1766	90	1676	24%	8%	12%
Investment receivables	17600	777	177	600	16823	1960	14863	**77%**	23%	36%
Employment contract	11600	152	31	121	11448	1045	10403	**80%**	10%	18%
Information disclosure	1600	30	7	23	1570	164	1406	77%	12%	21%
Stock claims	14200	529	82	447	13671	1439	12232	**84%**	24%	37%
Moving trouble	5200	79	58	21	5121	715	4406	27%	3%	5%
Building surrender	1600	4	0	4	1596	88	1508	100%	4%	8%
Facility admission fee	5800	2	0	2	5798	360	5438	100%	1%	1%
Contract receivables	3800	31	15	16	3769	328	3441	52%	5%	9%
Road maintenance guarantee	8600	34	8	26	8566	516	8050	76%	5%	9%

$$\text{Precision} = \frac{TP}{TP+FP} \tag{8}$$

$$\text{F1} = \frac{2Recall \times Precision}{Recall + Precision} \tag{9}$$

TP: True Positive (i.e., when the actual word and predicted word are positive)

TN is a true negative (i.e., when the actual word and the predicted word are negative). FN is a false negative (i.e., when the actual word is positive, but the predicted word is negative).

FP is a false positive (i.e., when the actual word is negative, but the predicted word is positive). The actual positive is "TP + FN," and the total predicted positive is "TP + FP." "Recall" is the index that indicates the fraction that was correctly predicted among all the words that were positive. "Precision" is the index that shows the fraction of positive words among all the words that were predicted to be positive. F1 is the index that shows the balance of recall and precision. The results of the experiment in various types of anonymized precedents are shown in Table V.

If the CW_PPL of the confidential word ("A") was lower than 60 (the threshold), we defined it as the correct word (TP). If not, it would not be recognized (FN). As a result, the recall score showed good values for some types of precedents. In particular, when similar phrases appeared multiple times in some precedents, our proposed model could recognize the confidential word well. Also, if we had more total words in one precedent, the recall score was good ranging from 69% to 84%.

The confidential words were sometimes replaced with non-English letters, such as the Greek letters γ or β; then, our proposed model could not predict them as confidential words. However, if they are replaced as "A" during preprocessing, the recall score would be good. Fig. 11 shows how our proposed model anonymized the confidential word for the original non-anonymized precedents.

All the experiments described so far used test data that included the confidential words converted to "A." Our final goal is to automatically replace the anonymous words in the actual precedents to the capital letter "A" by using our proposed model. Therefore, we used the test data as the original precedent in which the confidential words were not converted to "A"; we experimented with this model to examine whether the model was practical or not.

B. *Experiment with Proposed Model Using Non-Anonymized Precedents*

In this case, we defined that if the CW_PPL score was less than 100 (the threshold), our proposed model was correctly predicted as the confidential word. We experimented with our proposed model in some actual non-anonymized precedents, as shown in Table VI.

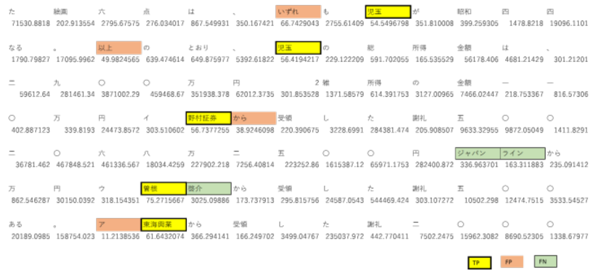

Fig. 11. Example of practical use in the Japanese precedent

In the case of the "Lockheed case," the recall was 62%, and the precision was 12%. One reason why the recall score was low was because our proposed model could not predict the name in an address correctly as a confidential word, especially because the addresses in the precedents were often very long. Another reason was that MeCab sometimes could not extract a person's name correctly when the full name appeared. In the case of "Traffic accidents," the recall score was 76%, which showed that our model was effective for practical use. In this case, if a person's name appeared in the precedent, our model could predict them correctly.

From these results, we can conclude that our model is not practical enough at present. However, if many confidential words appeared in the precedents, our proposed model was able to predict the confidential words correctly because the same phrases appeared in the Japanese precedents. Furthermore, if MeCab could accurately extract the personal name and another proper noun, our proposed model could prevent the false detection of confidential words that were not true. For the instances when our model could not predict long address that included numbers, such as 103-123–24-55, it will be effective to replace the long numbers with the letter "A" as the confidential word during preprocessing. Then, it would be possible to increase the possibility of practical use.

VI. CONCLUSION

Privacy is one of the most critical concerns in communication. If we can automatically hide confidential words, the communication becomes safe. Therefore, we developed a technology to predict confidential words that would use target data from Japanese judicial precedents.

However, it is not easy to completely anonymize confidential words such as personal names and locations. NER is probably the first step for information extraction; it seeks to locate and

TABLE VI. RESULT OF THE EXPERIMENT IN ACTUAL PRECEDENTS

Item	Total word	Confidential word			Normal word			Recall	Precision	F1
		Appear (TP+FN)	No-hit (FN)	Hit (TP)	Appear (TN)	Hit (FP)	No-hit			
Lockheed case	10000	286	109	177	9714	1297	8417	**62%**	**12%**	**20%**
Traffic accident	6000	189	46	143	5811	547	5264	**76%**	**21%**	**33%**

classify the named entities in Japanese precedents into predefined categories, such as the names of persons and locations. Nevertheless, the problem is that the cost of manually making such a corpus is very high.

One technique to protect privacy is to find confidential words in a file or on a website and convert them into meaningless words. Researchers have used the NER technique to extract the anonymous word. However, they need to prepare a proper noun dictionary that is updated to the latest issue; this is an expensive affair.

However, a neural network is useful for predicting the confidential words, and it is intelligent enough to anonymize confidential words automatically. Using Japanese judicial precedents, we have already proposed a recognition technique for confidential words using a neural network.

First, we proposed the Bi-directional LSTM–LR model that was effective for detecting the target words in long sequential words. However, the accuracy for detecting the confidential words (CW_PPL) in the precedents was poor [19]. To improve the CW_PPL score, we attempted to add additional information to the neural network called the POS tag. Two types of information (words and POS tags) needed to be entered into the neural network for learning. Then, we proposed a new model in which the Bi-directional LSTM–LR model was combined with the POS tag extracted by MeCab and a Japanese morphological analyzer.

Further, we reviewed the algorithm of preprocessing. We found that many confidential words replaced by a wide-width letter had been missed in the previous algorithm. Therefore, we improved upon the algorithm and experimented by using the new model. We found that the CW_PPL score had improved substantially. We obtained an accuracy improvement of 88% for detecting the confidential word (CW_PPL) and a 21% improvement for detecting the target word (PPL) as compared with the previously proposed model (Bi-directional LSTM-LR). Then, we evaluated our proposed model with the evaluation index, such as “recall” or “precision,” to determine whether our proposed model could have practical applications.

As a result, we confirmed that our proposed model could predict the confidential word for practical applications in some precedents, and the recall score improved from 69% to 84%. However, the overall results were not satisfactory. We experimented with the non-anonymized precedents to investigate the model for practical use. Our model is not good enough for practical use at present. If this model works well by making improvements (i.e., improvements in preprocessing and upgrading the MeCab), we will establish an automatic detector tool for confidential words in the future. Our proposed model will be a low-cost tool for detecting confidential words; therefore, it will be a valuable contribution in cyber courts.

ACKNOWLEDGMENT

We express special thanks to all the people who joined the Cyber Court Project. Also, we thank Prof. Baranyi to introduce the idea of CogInfoCom to us.

REFERENCES

[1] Baranyi, P., Csapo, A., “Definition and synergies of cognitive infocommunications,” Acta Polytechnica Hungarica, vol. 9 no. 1, pp. 67–83, 2012. DOI: 10.1109/CogInfoCom.2012.6422001

[2] Baranyi, P., Csapo, A., Sallai, G, “Cognitive Infocommunications (CogInfoCom),” Springer International, ISBN 978-3-319-19607-7, 2015.

[3] Graham Wilcock, Kristiina Jokinen, (2017). Bringing Cognitive Infocommunications to small language communities; Cognitive Infocommunications (CogInfoCom 2017) DOI: 10.1109/CogInfoCom.2017.8268253

[4] Willam&Mary Law School, “Center For Legal and Court Technology”, [online]. Available: http://law.wm.edu/academics/intellectuallife/researchcenters/clct/ Sep 5, 20170.8

[5] Takehiko, K., “Cybercourt - Court Technology”, Journal of the Japanese Society for Artificial Intelligence, vol. 23, no. 4, pp. 513–520, Jul. 2008

[6] Takehiko, K., “Court technology for civil procedure”, Festschrift for the seventieth birthday of Prof. Takeshi Kojima II, pp. 961, Sep. 2009

[7] NHK, “NHK Special” [online]. Available: http://www6.nhk.or.jp/special/detail/index.html?aid=20050213 Feb. 2, 2005

[8] MIC, "A study of applying ICT for legal service", Report of research funded by the Ministry of Internal Affairs and Communications in Japan in 2009, March. 2010

[9] Cabinet Office, [online]. Available: http://www5.cao.go.jp/keizaishimon/kaigi/minutes/2017/0609/shiryo_07.pdf

[10] LAI, Siwei, “How to generate a good word embedding”, IEEE Intelligent Systems, vol. 31, no. 6, pp. 5–14, 2016. DOI: 10.1109/MIS.2016.45

[11] Kiryu, Y., Ito, A., and Kanazawa, M., “Acta Polytechnica Hungarica”, Sep. 9, 2018 DOI: 10.12700/APH.16.2.2019.2.8

[12] Courts in JAPAN, [online]. Available: http://www.courts.go.jp/app/hanrei_jp/search1, Sep. 5, 2017

[13] Github. http://taku910.github.io/mecab/ Mar. 9, 2018

[14] Jing Li, Aixin Sun, Jianglei Han, Chenliang Li. “A Survey on Deep Learning for Named Entity Recognition”, cs.CL, 2018. arXiv:1812.09449

[15] Shotaro Misawa, Motoki Taniguchi, Yasuhide Miura, and Tomoko Ohkuma. Fuji Xerox Co. Ltd. “Character-based Bidirectional-LSTM CRF with words and characters for Japanese Named Entity Recognition Proceedings of the First Workshop on subword and Character Level Models in NLP”, 97102, 2017 DOI: 10.18653/v1/W17-4114

[16] Yamazaki M., Morita H., Komiya K., Kotani Y. “Extracting the Translation of Anime Titles from Web Corpora Using CRF”, JCKBSE 2014: Knowledge-Based Software Engineering pp.311-320 DOI: 10.1007/978-3-319-11854-3_26

[17] Houssen Gasmi, Jannik Laval, Abdelaziz Bouras. “Information Extraction of Cybersecurity Concepts: An LSTM Approach” Applied Sciences 9(19):3945 Sep.2019. DOI: 10.3390/app9193945

[18] TKC Co., [online]. Available: http://www.tkc.jp/law/lawlibrary, Sep. 6, 2017

[19] Kanazawa, M., Ito, A., Kiryu, Y., et al. “Improve accuracy of predicting confidential words in judicial precedents,” Technical Committee on Thought and Language July. 28, 2018

Appendix A

The algorithm of the Bi-directional LSTM combined POS tag

Input data

Input corpus : Japanese precedents

Preprocessing

Input data

$w_1^j = w_1\, w_2, \cdots\cdots w_j$ (no space)

*data cleaning # deleate punctuation mark and unnecessary words

*insert space between word (w_1)and word (w_2)

*extract the POS tag of the word

Output data

$w_1^j = w_1\, , w_2, \cdots\cdots, w_j$ (with space) : *word*

$p_1^j = p_1\, , p_2,\ \cdots\cdots,\ p_j$: *POS*

*p_1 is a POS corresponding to w_1

Main process

Input data

/parameter

hidden_layer_size = 100
batch_size = 200
chunk_size = 10
epochs = 100
learning_rate = 0.001
forget_bias = 1.0

/Word

Backword data $(w_i)_i^{bj}$: $w_1\, , w_2, \cdots w_{10}$

Forword data $(w_i)_i^{fj}$: $w_{-1}\, , w_{-2}, \cdots w_{-10}$

/POS

Backword POS data $(p)_i^{bj}$: $p_1\, , p_2, \cdots p_{10}$

Forword POS data $(p_i)_i^{fj}$: $p_{-1}\, , p_{-2}, \cdots p_{-10}$

#chunk size (window size) =10

Step1→ Initialize LSTM

Step2→ get the embedding vector for input data

for j in range (chunk_size)

$\boldsymbol{X_b} = (w_b)_{i=1}^j$: *Word(backword)*

$\boldsymbol{X_f} = (w_f)_{i=1}^j$: *Word(forword)*

$\boldsymbol{Y_b} = (p_b)_{i=1}^j$: *POS(backword)*

$\boldsymbol{Y} = (p_f)_{i=1}^j$: *POS(forword)*

Step3→ Ready lstm cell (tensorflow)

bw_ lstm = BasicLSTMCell ($\boldsymbol{X_b}$) :word (backword)

fw_ lstm = BasicLSTMCell ($\boldsymbol{X_f}$) :word(forword)

p_bw_ lstm = BasicLSTMCell ($\boldsymbol{Y_b}$) :POS (backword)

p_fw_ lstm = BasicLSTMCell ($\boldsymbol{Y_f}$) :POS(forword)

:backword_outputs $h_t^{(b)} = Wh_t + Wh_{t+1} + b$

:forword_outputs $h_t^{(f)} = Wh_t + Wh_{t-1} + b$

:POS_backword_outputs $s_t^{(b)} = Ws_t + Ws_{t+1} + b$

:POS forword outputs $s_t^{(f)} = Ws_t + Ws_{t-1} + b$

**W*: weight, *b*:bias

Output

#the last output is the model's output

outputs=concat(bw_outputs,af_outputs,H_bw_outputs,H_fw_outputs)

outputs $\boldsymbol{Y} = (h_t^{(b)} + h_t^{(f)} + s_t^{(b)} + s_t^{(f)})$

output= Y [outputs] $\times W + b$:loss

$Y = (y_i)_i^j = \left(y_1, y_2, \cdots\cdots\cdots, y_j\right)$: *word*

Masakazu Kanazawa graduated Oyama Technical College in 1972. He developed the international digital telephone exchange system at KDD co. from 1972 to 2002. He joined KYOWA EXEO co., engaged in mobile system construction from 2002 to 2019. He received B.S from University of Human Arts and Sciences in 2007. He is a student of Graduate School of Engineering of Utsunomiya University now.

Atsushi Ito received B.S. and M.S. degrees from Nagoya University in 1981 and 1983, respectively. He also received a Ph.D. degree from Hiroshima City University in 2007. From 1983 to 2014, he was with Research and Development Laboratories of KDDI Corporation. He is now a Professor of the Graduate School of Engineering of Utsunomiya University from 2014. During 1991-1992, he was a visiting scholar at the Center for the Study of Language and Information (CSLI) of Stanford University. His current research interests include an open platform for mobile communications, ad hoc networking, user interface design, and ICT based agriculture.

Yuya Kiryu received B.S. and M.S. degrees from Utsunomiya University in 2016 and 2018, respectively. He entered in KDDI Corporation in 2018, and he engaged in the investment in mobile base station construction utilizing big data.

Kazuyuki Yamasawa was received MBA degree from Nihon University Graduate School of Global Business. He Joined TKC in 1986, engaged in system development for 2 years and he engaged in sales, finance and salary system proposal sales for more than 500 companies since 1988, making a top sales. He was appointed Managing Executive Officer in 2019, in charge of the legal information service business.

Takehiko Kasahara received Ph.D. in Civil Law, Graduate School of Law, Keio University in March 1988. He was Assistant of Faculty of Law, Saarland University, Germany in 1984. He is a Professor of Toin Yokohama University Graduate School of Law from 2003 to now. He is a member of the Private University Information Education Association Legal Education FD / IT Utilization Research Committee (1999-2018), Fellow of Nagoya University Law and Information Research Center (JaLII) (2008-), a Director of Information Network Law Society Research, Researcher of Research on ICT utilization promotion in legal services" (2009-2010) of Ministry of Internal Affairs and Communications ICT Advanced Business International Expansion Project ICT Utilization Rule Promotion Project (Cyber Special Zone. And he is a JURISIN steering member.

Fubito Toyama is an associate professor at the Faculty of Engineering, Utsunomiya University. He received his Ph.D degrees in engineering from Utsunomiya University in 2000. His current research interests are evolutionary computation, combinatorial optimization and bioinformatics. He is a member of IPSJ, IEICE and ITE.

Answer to the reviewer's comments

1. Please make sure that for all those articles that are cited, and has an existing DOI (at CrossCiteRef), the DOI is actually referenced in your paper.
→ OK! We added the DOI in reference paper (DOI included)

2.1 The main contribution must be specified.
→ We added our contribution in abstract

2.2 The motivation of the approach based on neural networks and the treated architectures must also be specified in the context of other similar modeling approaches. Authors' strong past papers in the field ought to be mentioned in this discussion.
→ we described in conclusion as additional sentence.

2.3 Discussion of related work on other nonlinear modeling approaches should be extended with the following papers, which recently came into my attention because they proved to be successful in various applications:
→ Thank you for presenting a reference paper. We think so that these technique is excellent, however, the cost of our proposed model is not high compared to these technique.

2.4 The paper contains some grammar problems. Their correction is needed
→ Thank you for your check. We revised our paper.

2.5 Please organize the modeling approach as an identification algorithm with clear steps. What about the convergence of the identification algorithm? Please discuss.
→ We described the identification algorithm at end of the paper body.

2.6 How is the network trained? Please present details on that in relation with the network architecture and the comment 5).
→ This is also described in the identification algorithm.

2.7 It is also not clear how the testing and validation are done.
→ This is also described in the identification algorithm.

2.8 The very good accuracy might indicate overfitting. Please comment.
→ We input the various kinds of precedents randomly to prevent overfitting. However, we think it is necessary to continuing more actual precedents. If overfitting would occur, we adopt the dropout layer
.

2.9 The validation would be better supported if you would add a link to the programs and datasets. In other words, you could save the programs and datasets in a webpage/repository and next cite the link to that webpage or repository in the paper body. This will ensure a sound validation, which is important in neural network and modeling approaches. This will also solve a part of the above and below comments
→ OK！I add the program in repository in the paper body.

2.10 This is an application paper. It is not clear how the theory from the previous sections is applied in Section V. More details are necessary. The comment 9) helps.
→ OK！

2.11 Transitions from section to section should be smoother. The comments 3) and 9) help.
→ OK！I add the program in repository in the paper body.

２.1２ The number of references is rather low for a strong journal paper. The comment 3) helps.
→ Thank you for your suggestion.

2.1３ Do you have random parameters in the algorithms? If yes, which are their effects?
→ Yes. We experimented changing the various parameter. However, we had get no effective results.

2.1４ Do you have comparisons with other modeling approaches including neural network-based ones? Please discuss
→ No, there are some paper regarding with the detecting the confidential word using named extraction method (using dictionary) but I couldn't find paper using the neural network only for detecting the confidential word.