Self-supervised Mining of Human Activity from CGM

Nguyen Minh The, Takahiro Kawamura, Hiroyuki Nakagawa, Yasuyuki Tahara, and Akihiko Ohsuga

Graduate School of Information Systems, The University of Electro-Communications, Tokyo, Japan
{minh,kawamura,nakagawa,tahara,akihiko}@ohsuga.is.uec.ac.jp

Abstract. The goal of this paper is to describe a method to automatically extract *all* basic attributes namely *actor*, *action*, *object*, *time* and *location* which belong to an activity, and the *transition* between activities in each sentence retrieved from Japanese CGM (consumer generated media). Previous work had some limitations, such as high setup cost, inability of extracting all attributes, limitation on the types of sentences that can be handled, and insufficient consideration of interdependency among attributes. To resolve these problems, this paper proposes a novel approach that treats the activity extraction as a sequence labeling problem, and automatically makes its own training data. This approach has advantages such as *domain-independence*, *scalability*, and *unnecessary hand-tagged data*. Since it is unnecessary to fix the positions and the number of the attributes in activity sentences, this approach can extract *all* attributes and transitions between activities by making *only a single pass* over its corpus.

1 Introduction

The ability of computers to provide the most suitable information based on users' behaviors is now an important issue in context-aware computing [1], ubiquitous computing [11] and social computing [12]. For example, a service delivers shop information based on the users' next destination [24], a service displays advertisements based on the users' behaviors [16], or an experience-sharing service as shown in Figure 1.

To identify the users' behaviors, it is necessary to understand *how to collect activity data*, *how to express or define each activity* and *its relationships*. It is not practical to define each activity and its relationships in advance, because it not only takes enormous cost, but also cannot deal with unpredictable behaviors. On the other hand, there are some projects that are collecting users' everyday event logs by using sensors installed in mobile phone, GPS or RFID tag [2,3], such as the *My Life Assist Service* [24], and the *Life Log* [4]. From these event logs, they try to extract users' activities, and predict the relationships between them. However, Kawamura et al. [10] indicated some problems in this approach, such as large amount of noisy data in event logs, high computational cost, security and privacy

B.-H. Kang and D. Richards (Eds.): PKAW 2010, LNAI 6232, pp. 61–72, 2010.

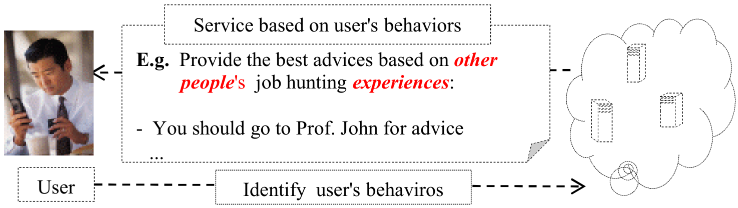

Fig. 1. Experience sharing service

issues. Additionally, if we do not have enough event logs of a large number of different users, it is difficult to discover common or exceptional action patterns.

With such problems discussed above, we try another approach that collects human activity data from CGM. Today, CGM is generated by users posting their activities to Twitter, Facebook, their weblogs or other social media. Thus, it is not difficult to collect activity data of different users from CGM. However, sentences retrieved from CGM have various structures, are complex, are syntactically incorrect. Thus, there are lots of challenges to extract all activity attributes and transitions between activities in these sentences. Few previous works have tried to extract attributes in each sentence retrieved from CGM. These works have some limitations, such as high setup costs because of requiring ontology for each domain [10]. Due to the difficulty of creating suitable patterns, these works are unable to extract all attributes [5,10], limited on the types of sentences that can be handled [5,6], and insufficiently consider interdependency among attributes [5,6].

Since each attribute has interdependent relationships with the other attributes in every activity sentence, we can treat attribute extraction as an *open relation extraction* [13]. In other words, we extract an action and other word phrases that have relationships with this action and describe their activity. In this paper, we propose a novel approach based on the idea of O-CRF [9] that applies self-supervised learning (Self-SL) and uses conditional random fields (CRFs) to the open relation extraction. O-CRF is the state-of-the-art of the open relation extraction from English web pages. Our approach focuses on Japanese CGM, and treats activity extraction as a sequence labeling problem. This approach automatically makes its own training data, and uses CRFs as a learning model. Our proposed architecture consists of two modules: Self-Supervised Learner and Activity Extractor. Given some activity sentences retrieved from the "people" category of Wikipedia, the Learner extracts all attributes and transitions between activities by using deep linguistic parser, and then automatically makes training data. The Learner uses CRFs to make the feature model of these training data. Based on this feature model, the Extractor automatically extracts all attributes and transitions between activities in each sentence retrieved from Japanese CGM.

The main contributions of our approach are summarized as follows:

– It is *domain-independent*, without requiring *any* hand-tagged data.
– It can extract *all* attributes and transitions between activities by making only a *single pass* over its corpus.

– It can handle *all* of the standard sentences in Japanese, and achieves high precision on these sentences.
– It can avoid the privacy problem.

The remainder of this paper is organized as follows. In section 2, we indicate challenges of extracting attributes in more detail. Section 3 explains how our approach makes its own training data, and extracts activity in each sentence retrieved from Japanese CGM. Section 4 reports our experimental results, and discuss how our approach addresses each of the challenges to extract activity attributes. Section 5 considers related work. Section 6 consists of conclusions and some discussions of future work.

2 Challenges

2.1 Activity Attributes Definition

The key elements of an activity are actor, action, and object. To provide suitable information to users, it is important to know *where and when activity happens*. Therefore, in this paper, we define an activity by five basic attributes: actor, action, object, time, and location. We label these attributes as *Who*, *Action*, *What*, *When* and *Where* respectively, and label the transitions between activities as *Next* or *After*. For example, Figure 2 shows the attributes and the transition between activities derived from a Japanese sentence.

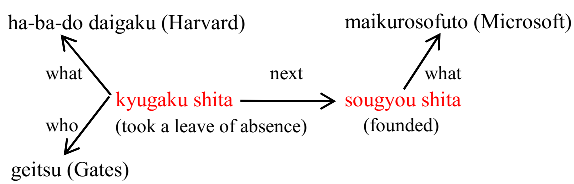

Fig. 2. The attributes and the transition between the activities derived from the activity sentence "*geitsu* ha *ha-ba-do daigaku* wo *kyugaku shi*, *maikurosofuto* wo *sougyou shita*" (Gates took a leave of absence from Harvard, then founded Microsoft).

2.2 Challenges of Extracting Activity Attributes

Extracting activity attributes in sentences retrieved from CGM has many challenges, especially in Japanese. Below, we explain some of them:

1. As shown in Figure 3, O-CRF extracts binary relations in English, and these relations must occur between entity's names within the same sentence [9]. Japanese sentences do not follow this rule, thus we can not directly apply O-CRF for extracting activity attributes in Japanese.
2. In Japanese, there are not word spaces, further word boundaries are not clear. However, previous works in CRFs assume that observation sequence (word) boundaries were fixed. Therefore, a straightforward application of CRFs is impossible.

Fig. 3. Limitation of O-CRF

3. Whereas almost typical sentences in English follow the *subject-verb-object* rule [9], Japanese sentences are flexible with many types of structures.
4. Since number and position of attributes are changing in different sentences, it is difficult to create instances or patterns to extract all attributes and transitions between activities.
5. It is not practical to deploy deep linguistic parsers, because of the diversity and the size of the Web corpus [9]. Additionally, sentences retrieved from CGM are often diversified, complex, syntax wrong, and have emoticons. Therefore, the deep linguistic parsers often have errors when parsing these sentences.
6. If extraction method is domain-dependent, then when shifting to a new domain it will require a new specified training examples. And, the extraction process has to be run, and re-run for each domain.

3 Self-supervised Mining of Human Activity

3.1 Activity Extraction with CRFs

CRFs [17] are undirected graphical models for predicting a label sequence to an observed sequence. The idea is to define a conditional probability distribution over label sequences given an observed sequence, rather than a joint distribution over both label and observed sequences. CRFs offers several advantages over hidden Markov models and stochastic grammars, including the ability of relaxing strong independence assumptions made in those models. Additionally, CRFs also avoids the label bias problem, which is a weakness exhibited by maximum entropy Markov models (MEMMs) and other conditional Markov models based on directed graphical models. CRFs achieves high precision on many tasks including text chunking [18], named entity recognition [19], Japanese morphological analysis [20].

By making a first-order Markov assumption that has dependencies between output variables, and arranging variables sequentially in a linear chain, activity extraction can be treated as a sequence labeling problem. Figure 4 shows an example where activity extraction is treated as a sequence labeling problem. Tokens in the surrounding context are labeled using the IOB2 format. B-X means "begin a phrase of type X", I-X means "inside a phrase of type X" and O means "not in a phrase". IOB2 format is widely used for natural language tasks. In this paper, we use CRF++[1] to implement this linear chain CRF.

[1] Available at http://crfpp.sourceforge.net/

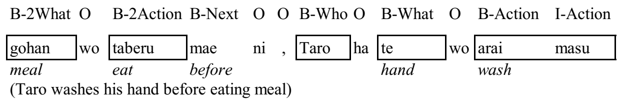

Fig. 4. Activity Extraction as Sequence Labeling

3.2 Proposed Architecture

As shown in Figure 5, the architecture consists of two modules: *Self-Supervised Learner* (I in Figure 5) and *Activity Extractor* (II in Figure 5). Sentences retrieved from the "people" category of Wikipedia are often syntactically correct, activity describable, and easy to parse. Therefore, we parse these sentences to get activity sentences (that describe activities), and then send these activity sentences as sample data to the Learner. The Learner deploys deep linguistic parser to analyze the dependencies between word phrases. Based on the prepared list of Japanese syntax, it selects trustworthy attributes to make training data, and the feature model of these data. The Extractor does *not* deploy deep parser, it bases on this feature model to automatically extract all attributes, and transitions between activities in sentences retrieved from Japanese CGM. Below, we describe each module in more detail.

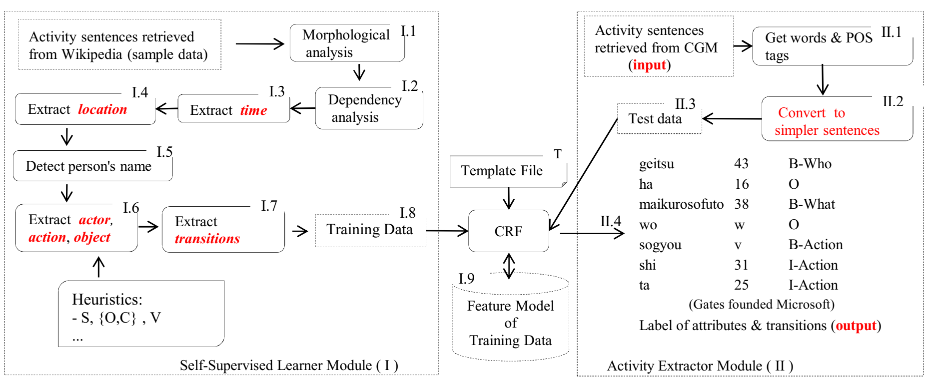

Fig. 5. Proposed Architecture: By using deep linguistic parser, the Learner automatically makes its own training data

3.2.1 Self-supervised Learner Module

We will use the example sentence "geitsu ha maikurosofuto wo sougyou shita" (Gates founded Microsoft) to explain how the Learner works and makes its own training data. As shown in Figure 5, the Learner consists of nine key tasks:

1. By using Mecab[2], it parses the sample data to get words and their POS tags in each sentence (I.1 in Figure 5).

[2] Available at `http://mecab.sourceforge.net/`

2. By using Cabocha[3], it analyzes the interdependencies among word phrases in each sentence (I.2 in Figure 5). Up to this step, the Learner can detect verb phrase (VP), noun phrase (NP), POS tags, named entity, and the interdependencies among word phrases in each sentence.
3. In addition to the above analytical result, based on the Japanese regular time-expressions such as VP-taato, VP-maeni, toki...etc, the Learner extracts the time of activity and labels it as *When* (I.3 in Figure 5).
4. To improve precision of location extraction, in addition to the above analytical result, the Learner uses the Google map API to extract the location of activity and labels it as *Where* (I.4 in Figure 5).
5. Japanese natural language processing (NLP) tools often have errors when analyzing foreign person name. In this case, the Learner utilizes the "human names" category of Wikipedia to improve precision of person name detection (I.5 in Figure 5).
6. To select trustworthy activity sentences, we prepare a list of all Japanese syntax patterns such as "S, {O, C}, V", "{O, C}, V, S"...etc. Where S means subject, O means object, C means complement, V means verb. Actor, action, object correspond to S, V, O respectively. Based on these syntax patterns, the Learner extracts actor, action, object, and then labels them as *Who*, *Action*, *What* respectively (I.6 in Figure 5).
7. Based on syntax patterns such as V-taato...etc, the Learner extracts the transitions between activities, and labels as *Next* or *After* (I.7 in Figure 5).
8. As shown in Figure 6, training data are automatically created by combining the above results (I.8 in Figure 5).

B-Who	O	B-What	O	B-Action	I-Action
43	16	45	w	v	25
geitsu	ha	maikurosofuto	wo	sougyou	shita
Gates		Microsoft		founded	(Gates founded Microsoft)

Fig. 6. Training data of the example sentence

9. The Leaner uses CRF and template file to automatically generate a set of feature functions (f 1, f 2, ..., f n) as illustrated in Figure 7. The feature model of these training data is created from this set of feature functions (I.9 in Figure 5).

```
f 1 = if (label = "B-Who" and POS="43") return 1 else return 0
....
f n = if (label = "B-Action" and POS="v") return 1 else return 0
```

Fig. 7. Feature functions

[3] Available at `http://chasen.org/~taku/software/cabocha/`

3.2.2 Activity Extractor Module

We parse Japanese CGM pages to receive activity sentences, and then remove emoticons, and stop words in these sentences. In this case, stop words are the words which do not contain important significance to be used in activity extraction. After this pre-processing, we send activity sentences to the Extractor. As shown in Figure 5, the Extractor consists of four key tasks:

1. The Extractor uses Mecab to get words and their POS tags (II.1 in Figure 5). As shown in Figure 8, in addition to analytical result by Mecab, the Extractor utilizes html tags to detect a long or complex noun phrases.

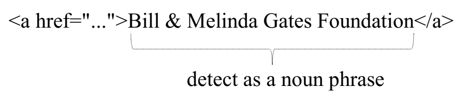

Fig. 8. Using html tags to detect a noun phrase

2. Sentences retrieved from CGM are complex, thus the Extractor converts these sentences to simpler sentences by simplifying noun phrases and verb phrases (II.2 in Figure 5). When converting, it keeps the POS tags of these word phrases.
3. The Extractor makes test data by combining the above results (II.3 in Figure 5). As shown in Figure 9, unlike training data, test data does not have label row. This label row is predicted when testing.

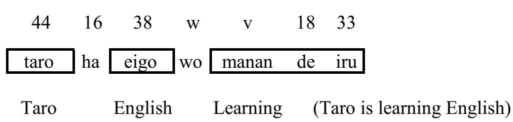

Fig. 9. Test data for the activity sentence *"taro ha eigo wo manan de iru"*

4. Based on the feature model, the Extractor automatically extracts all attributes and transitions between activities in each sentence of the test data (II.4 in Figure 5).

3.2.3 Template File

We use the feature template file to describe features that are used in training and testing (T in Figure 5). The set of features includes words, verbs, part-of-speech (POS) tags and postpositional particles in Japanese. To model long-distance relationships, this paper uses a window size of 7.

4 Evaluation

4.1 Experimental Results

To evaluate the benefits of our approach, we used the set of 533 activity sentences[4] randomly retrieved from Japanese CGM. There are 356 sentences that describe one activity, 177 sentences that describe two activities in this experimental data. Figure 10 shows two sentences which are used for this experiment.

kaunta- de, nihon no menkyosyou wo teiji shite tetsudzuki wo okonau
counter Japanese driver's license show then procedure do
(At the counter, shows the Japanese driver's license and then proceeds)

heya he modo te gaisyutsu no jyunbi wo shimashita
room come back going out preparation done
(Came back to the room, then prepared to go out)

Fig. 10. Two activity sentences in our experimental data

In this experiment, we say an activity extraction is correct when all attributes of this activity are correctly extracted. The precision of each attribute is defined as the number of correctly extracted attributes divided by the total number. Using one PC (CPU: 3.2GHz, RAM: 3.5GB), the Extractor module makes only a single pass over the entire experimental data set, and gets the results[5] as shown in Table 1. This process took only 0.27s.

Table 1. Experimental Results

@	Should be extracted	Correct	Precision (%)
Activity	710	631	88.87
Actor	196	182	92.86
Action	710	693	97.61
Object	509	479	94.11
Time	173	165	95.38
Location	130	120	92.31
Transition	26	22	84.62

4.2 Consideration

The experimental results shown that our approach can automatically extract *all* attributes and transitions between activities in each sentences by making *only a single pass* with high precision. Additionally, our method took only 0.27s, while a widely known deep parser such as Cabocha took over 46.45s for parsing the experimental data (our approach outperforms over 172 times). We consider the experimental results as follows:

[4] Available at http://docs.google.com/View?id=dftc9r33_1077g63vrjc5
[5] Available at http://docs.google.com/View?id=dftc9r33_1078cr9hd3mt

– In activity sentence, action corresponds to a verb phrase. We used Mecab, the parser with high precision of detecting verb phrase. Additionally, we simplified complex verb phrases before testing. These are the reasons of high precision (97.61%) when extracting *action.*
– In Japanese sentence, postpositional particles (wo, ni, he) are often allocated between action and object. In addition to this special feature, we simplify complex noun phrases before testing. So that, our approach achieved high precision (94.11%) of *object* extraction.
– In this experimental data, many actors are people's name. Moreover, in typical Japanese sentence, there are "ha" or "ga" behind subject. Therefore, by using these features, we achieved high precision (92.86%) of *actor* extraction.
– In addition to parsing result of Mecab, we utilize expressions of time in Japanese. So that, our approach achieved high precision (95.38%) of *time* extraction.
– By using Google Map API, we can deal with long or complex addresses. This is reason for high precision (92.31%) of *location* extraction.
– In this experimental data, transitions between activities are explicit. So that, we can extract *transitions* between activities in this data.

Below, we describe how our approach resolves the limitations of the previous works outlined in introduction section.

– It is domain-independent, and automatically creates training data. Therefore, our approach does not take high setup costs.
– By treating activity extraction as a sequence labeling problem, and not fixing the position and number of attributes, our approach is able to extract all attributes in an activity sentence.
– We create training data for all typical sentences. Additionally, we simplify complex sentences before testing. These are reasons for which the Extractor could deal with many type of sentences.
– The feature model contains features of dependencies between attributes in each sentence of training data. Based on these features, the Extractor can consider dependencies between attributes in each sentence of testing data.
– From public CGM, we can collect activity data of many different users. And, this approach can avoid privacy problem.

Our approach addresses each of the challenge indicated in section 2 as follows:

1. It does not fix the position and number of attributes in activity sentences. Additionally, it uses the heuristics (syntax patterns) to select trustworthy training data. We also design the template file to handle multi-attributes in each activity sentence.
2. It uses Mecab and html tags to get word phrases in each sentence.
3. Based on a list of all Japanese syntax, it makes training data for all typical sentences.
4. It treat activity extraction as labeling problem. Additionally, it uses machine learning approach to inference new patterns.
5. The Extractor does not deploy deep linguistic parsers. It deletes emoticons and stop words in each sentences retrieved from CGM. Additionally, the Extractor simplifies complex sentences before testing.

6. The Leaner uses common syntax patterns which do not depend on a specified domain. In other words, the Leaner can make training data for any domain.

However, our approach also has some limitations. Firstly, it only extracts activities that are explicitly described in the sentences. Secondly, it has not yet extracted transitions between activities in document-level. Finally, to handle more complex or incorrect syntax sentences, we need improve our architecture.

4.3 Applying to Other Languages

Our proposed architecture focus on Japanese, but it could be applied to other languages by changing suitable syntax patterns (heuristics) for the Leaner. We should also re-design the template file to utilize special features of the applied language.

4.4 Comparison with O-CRF

Because of the differences in tasks (activity, binary relation) and languages (Japanese, English), it is difficult to compare our approach with O-CRF. We try to compare them according to the some criteria as shown in Table 2.

Table 2. Comparison with O-CRF

	O-CRF	Our Method
Language	English	Japanese
Target data	Binary relation	Human activity
Type of sentences can be handled	S-V-O	{O, C}, V; S, {O, C}, V;... all typical syntax
Relation must occur between entities	YES	NO
Requirement of determining entities before extracting	YES	NO

5 Related Work

There are two fields related to our research: relation extraction (RE) and human activity extraction (AE) from the Web. Below, we discuss the previous researches of each field.

5.1 Relation Extraction

The main researches of RE are DIPRE [14], SnowBall [15], KnowItAll [8], Pasca [7], TextRunner [13], O-CRF [9] (the upgraded version of TextRunner).

DIPRE, SnowBall, KnowItAll, and Pasca use bootstrapping techniques applied for unary or binary RE. Bootstrapping techniques often require a small set of hand-tagged seed instances or a few hand-crafted extraction patterns for each domain. In addition, when creating a new instance or pattern, they could possibly extract unwanted patterns around the instance to be extracted, which would lead to extract unwanted instance from the unwanted patterns. Moreover, it is difficult to create suitable instances or patterns for extracting the attributes and relationships between activities appeared in sentences retrieved from the Web.

TextRunner is the first Open RE system, it uses self-supervised learning and a Naive Bayes classifier to extract binary relation. Because this classifier predicts the label of a single variable, it is difficult to apply TextRunner to extract all of the basic attributes.

5.2 Human Activity Extraction

Previous works on this field are Perkowitz [5], Kawamura [10] and Kurashima [6]. Perkowitz's approach is a simple keyword matching, so it can only be applied for cases of recipe web pages (such as making tea or coffee). Kawamura's approach requires a product ontology and an action ontology for each domain. So, the precision of this approach depends on these ontologies.

Kurashima used JTAG [21] to deploy a deep linguistic parser to extract action and object. It can only handle a few types of sentences, and is not practical for the diversity and the size of the Web corpus. Additionally, because this approach gets date information from date of weblogs, so it is highly possible that extracted time might not be what activity sentences describe about.

6 Conclusions and Future Work

This paper proposed a novel approach that automatically makes its own training data, and uses CRFs to automatically extract *all* attributes and transitions between activities in each sentence retrieved from Japanese CGM. Without requiring any hand-tagged data, it achieved high precision by making only a *single pass* over its corpus. This paper also explained how our approach resolves the limitations of previous works, and addresses each of the challenges to activity extraction.

We are improving the architecture to handle more complex or syntax incorrect sentences. Based on links between web pages, we will try to extract transitions between activities at the document-level. In the next step, we will use a large data set to evaluate our approach. We are also planning to build a large human activity semantic network based on mining human experiences from the entire CGM corpus.

References

1. Matsuo, Y., Okazaki, N., Izumi, K., Nakamura, Y., Nishimura., T., Hasida, K.: Inferring long-term user properties based on users' location history. In: Proc. IJCAI 2007, pp. 2159–2165 (2007)
2. Pentney, W., Kautz, H., Philipose, M., Popescu, A., Wang, S.: Sensor-Based Understanding of Daily Life via Large-Scale Use of Common Sense. In: Proc. AAAI 2006 (2006)

3. Pentney, W., Philipose, M., Bilmes, J., Kautz, H.: Learning Large Scale Common Sense Models of Everyday Life. In: Proc. AAAI 2007 (2007)
4. KDDI, Corp.: My Life Assist Service (2009), `http://www.kddilabs.jp/english/tech/frontier.html`
5. Perkowitz, M., Philipose, M., Fishkin, K., Patterson, D.J.: Mining Models of Human Activities from the Web. In: Proc. WWW 2004 (2004)
6. Kurashima, T., Fujimura, K., Okuda, H.: Discovering Association Rules on Experiences from Large-Scale Weblogs Entries. In: ECIR 2009, pp. 546–553 (2009)
7. Pasca, M., Lin, D., Bigham, J., Lifchits, A., Jain, A.: Organizing and Searching the World Wide Web of Facts - Step One: the One-Million Fact Extraction Challenge. In: Proc. AAAI 2006, pp. 1400–1405 (2006)
8. Etzioni, O., Cafarella, M., Downey, D., Kok, S., Popescu, A., Shaked, T., Soderland, S., Weld, D., Yates, A.: Methods for Domain-Independent Information Extraction from the Web: An Experimental Comparison. In: Proc. AAAI 2004 (2004)
9. Banko, M., Etzioni, O.: The Tradeoffs Between Open and Traditional Relation Extraction. In: Proc. ACL 2008 (2008)
10. Kawamura, T., Nguyen, M.T., Ohsuga, A.: Building of Human Activity Correlation Map from Weblogs. In: Proc. ICSOFT (2009)
11. Poslad, S.: Ubiquitous Computing Smart Devices, Environments and Interactions. Wiley, Chichester (2009), ISBN: 978-0-470-03560-3
12. Ozok, A.A., Zaphiris, P.: Online Communities and Social Computing. In: Third International Conference, OCSC 2009, Held as Part of HCI International 2009, San Diego, CA, USA. Springer, Heidelberg (2009), ISBN-10: 3642027733
13. Banko, M., Cafarella, M.J., Soderland, S., Broadhead, M., Etzioni, O.: Open information extraction from the Web. In: Proc. IJCAI 2007, pp. 2670–2676 (2007)
14. Brin, S.: Extracting Patterns and Relations from the World Wide Web. In: WebDB Workshop at 6th International Conference on Extending Database Technology, EDBT 1998, Valencia, Spain, pp. 172–183 (1998)
15. Agichtein, E., Gravano, L.: Snowball: Extracting relations from large plain-text collections. In: Proc. ACM DL 2000 (2000)
16. Peppers, D., Rogers, M.: The One to One Future. Broadway Business (1996), ISBN-10: 0385485662
17. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields Probabilistic models for segmenting and labeling sequence data. In: Proc. ICML, pp. 282–289 (2001)
18. Sha, F., Pereira, F.: Shallow parsing with conditional random fields. In: Proc. HLTNAACL, pp. 213–220 (2003)
19. McCallum, A., Li, W.: Early results for named entity recognition with conditional random fields, feature induction and Web-enhanced lexicons. In: Proc. CoNLL (2003)
20. Kudo, T., Yamamoto, K., Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphologiaical Analysis. In: IPSJ SIG Notes, pp. 89–96 (2004)
21. Fuchi, T., Takagi, S.: Japanese morphological analyzer using word co-occurence-JTAG. In: Proc. ACL 1998, pp. 409–413 (1998)
22. Kudo, T., Matsumoto, Y.: Japanese Dependency Analysis using Cascaded Chunking. In: Proc. CoNLL 2002, pp. 63–69 (2002)
23. Hiroyuki, Y., Hideyuki, T., Hiromitsu, S.: An individual behavioral pattern to provide ubiquitous service in intelligent space. WSEAS Transactions on Systems, 562–569 (2007)
24. NTTDocomo, Inc.: My Life Assist Service (2009), `http://www.igvpj.jp/contents_en/activity09/ms09/list/personal/ntt-docomo-inc-1.html`