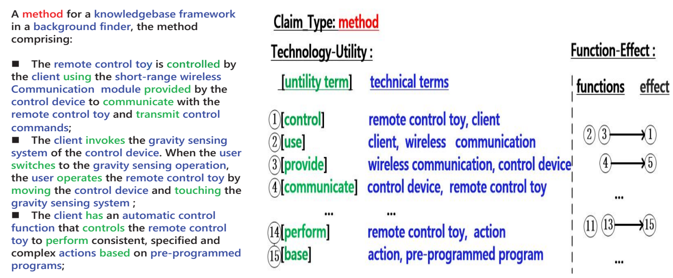
Patent Technical Function-effect Representation and Mining Method

Weidong Liu, Piying Zhang, Wenbo Qiao
College of Computer Science, Inner Mongolia University
Inner Mongolia Key laboratory of Social Computing and Data Processing
cslwd@imu.edu.cn, zhangpy.imu@gmail.com, imu.qiaowb@gmail.com

***Abstract*—With increasing global competition of intellectual property, a large number of unstructured patent texts are generated for technology protection. The ocean of patent texts include many long sentences about technologies, technical functions, technical effects and complexity relations between them, which make it difficult to textual representation and mining. To solve the above issues, we represent a patent by its technical function-effects, which are mined from the patent. The model represents functions/effects by valence utility-technologies and represents text by association relations between functions and effects. We evaluate our model by comparing with the state-of-the-art models on the patent data set. The results show that our model outperforms other models in evaluation measurement. Such representation can be applied to patent information retrieval and patent text analysis.**

***Index Terms*—patent text, representation, function-effect**



(a) Original Claims (b) Structured Claims

Fig. 1. Compared with original claim, the structured claim consist of some utility-technologies and function-effects.

I. Introduction

With the economic globalization, technological innovation are fueling the knowledge-economy increase. To realize innovation-driven economic development, it is urgent to create, protect and applicant for the patents.

Manually reading and comparing the ocean of patents is time-consuming, since the patents have many long sentences which include some technologies, technical functions, technical effects and some complex relations between them. As shown in Fig.1(a), the original claim of patent[1] have many long sentences, which consist of technical terms(marked by blue font) and utility terms(marked by green font). Some implicit relations are included in the patent as well, such as the utility-technology relation between 'control' and 'control device'and the function-effect relation between 'use: wireless communication', 'provide: control device' and 'control: toy client' in Fig.1(b).

Compared with the expected representation in Fig.1(b), most previous representation models lack some aspects of consideration to these technical functions/effects and their relations.

The current textual representation are summarized into three categories:

1) Concept-based models. The concept-based models often represent text by a set of concepts, which include explicit textual representation and implicit text representation. 1) In explicit textual representation, a concept, a topic or a text is represented by a vector, where each element corresponds to a clear semantic meanings. The most used models include vector space model(VSM)[1, 2], explicit semantic analysis model(ESA)[3], probability-based topic model, object attribute model(OAR)[4] etc. 2) Implicit textual representation models map the textual information to a latent vector space. The semantic information is often no-interpretable. The most used models include implicit semantic analysis model(LSA)[5], neuron-language model (NLM)[6, 7], segment vector model(PV)[8], hyperspace simulation language model(HAL)[8, 9] and word2vec[10, 11].
2) Relation-based models. Relation based models often represent text as a set of concepts and their relations. Shinmori represents patent claims by six types of relation(process, composition, characteristics, premise, combination)[12]. Okamoto represents patent claims by verb-nouns relations[13]. Luo represents text by association relations[14]. Besides, the relations can form some networks, such as semantic linking network[15], association linking network[16] and knowledge graphs[17–19].
3) Other models. Temporal, citation and other features are considered in textual representation[20–24].

When the above models are used in patent textual representation, the limitations are summarized as follows.

1) Overlooked technical functions/effects in the patent. Most previous models represent a patent by some technical terms(the nouns) other than the functions/effects of the patent.

[1]No.CN102800178A

DOI reference number: 10.18293/SEKE2020-034

2) Neglected the relations between functions and effects. Most models only focus on the association relations between technical terms other than the function-effects.

To overcome the above limitations, we propose a patent function-effect representation and its mining method, by which each the patent is represented by some function-effects. The remainders of this paper are organized as follows. Section 2 introduces the preliminary knowledge. Section 3 proposes a function-effect based patent representation. Section 4 gives a mining method of function-effect. Section 5 reports experiments. Section 6 makes conclusion.

II. Preliminary Knowledge

Valence is a term in chemistry, which refers to the ability of an atom to combine with other atoms. In linguistics, valence refers to the number of arguments(nouns) a verb carries. Inspired by the hypothesis of valence theory, each function/effect of the technology can be represented by the hypothesis of valence theory.

Definition 1: Hypothesis of Valence Theory(HVT)

$$HVT = \{VT^k | 0 \leq k < |HVT| - 1\} \tag{1}$$

$$VT^k = \{vt_i^k | 0 \leq i < |VT^k| - 1, |vt_i^k| = k\} \tag{2}$$

$$vt_i^k = v_{(i,0)}^k : t_{(1:|vt_i^k|)}^k \tag{3}$$

Where v denotes a utility term; t denotes a technical term; $VT^{(k)}$ is a set of k-$valenc$ utility-technologies; vt_i^k is the i^{th} a utility-technology in $VT^{(k)}$; $|vt_i^k|$ denotes the number of technical term in vt_i^k.

For the claims of the previous patent, the hypothesis of valence theory(HVT) is shown in TableI.

TABLE I
HYPOTHESIS OF VALENCE THEORY OF THE PATENT CLAIMS

$HVT(p)$	$HVT(p) = \{VT^1, VT^2, VT^3\}$
VT^1	$v_9 : t_4, \quad v_{10} : t_6, \quad v_2 : t_2 t_3, \quad v_3 : t_3 t_4, \quad v_4 : t_4 t_1$
VT^2	$v_5 : t_4 t_5, \quad , v_7 : t_7 t_8, \quad v_8 : t_7 t_1, \quad v_{11} : t_2 t_9, \quad v_1 : t_1 t_9, \quad v_{12} : t_1 t_{10}, \quad v_{13} : t_{10} t_{11}$
VT^3	$v_6 : t_2 t_4 t_6$
Technical Terms: t_1:remote control toy, t_2:client, $\ldots t_{11}$: pre-programmed program **Utility Terms**: v_1: control, v_2: use, v_3: provide, $\ldots v_{12}$: perform, v_{13}: base	

HVT can represent fucntions/effects of the patent by utility-technologies.

III. Function-effect based Patent Representation

To represent the functions, the effects and their relations, a patent function-effect representation model is proposed.

Definition 2: Function-effect Representation(FR)

$$FR = \{\Phi^k | 0 \leq k < |FR| - 1\} \tag{4}$$

$$\Phi^k = \{\phi_i^k | 0 \leq i < |\Phi^k| - 1, |\phi_i^k| = k\} \tag{5}$$

$$\phi_i^k = vt_{(i,1:|\phi_i^k|-1)}^k \rightarrow vt_{(i,0)}^k \tag{6}$$

where Φ^k denotes a set of k-$degree$ function-effects; ϕ_i^k denotes the i^{th} function-effect in Φ^k; $|\phi_i^k|$ denotes the degree of ϕ_i^k; $vt_{(i,j)}^k$ denotes the j^{th} utility-technology in ϕ_i^k.

TABLE II
FUNCTION-EFFECT REPRESENTATION OF THE PATENT CLAIMS

$FR(p)$	$FR(p) = \{\Phi^{(0)}, \Phi^{(2)}\}$
Φ^0	$\phi_1^0 = vt_4, \phi_2^{(0} = vt_5, \phi_3^0 = vt_6, \phi_4^0 = vt_7, \phi_5^0 = vt_{13}$
$\Phi^{(2)}$	$\phi_1^2 = vt_2 vt_3 \rightarrow vt_1, \phi_2^2 = vt_9 vt_{10} \rightarrow vt_8, \phi_3^2 = vt_{11} vt_1 \rightarrow vt_{12}$
valence utility-technologies: $vt \in HVT$ in tableI	

The function-effect representation(FR) of the previous patent claims is shown in TableII.

For the previous patent, its function-effect representation is shown in TableII.

TABLE III
SYMBOLS AND DESCRIPTION

Symbols	Description
$V = \{v_i \| 0 \leq i < \|V\| - 1\}$	a set of utility terms
$T = \{t_i \| 0 \leq i < \|T\| - 1\}$	a set of technical terms
$lc_i^k = t_{(i,1:\|lc_i^k\|-1)}^k \rightarrow t_{(i,0)}^k$	a association relation between technical terms
$\|lc_i^{(k)}\|$	the degree of $lc_i^{(k)}$
$LC^k = \{lc_i^k \| 0 \leq i < \|LC^k\| - 1\}$	a set of k-$degree$ relations
$HLC = \{LC^k \| 0 \leq k < \|HLC\| - 1\}$	hypothesis of linear concept
$vt_i^k = v_{(i,0)}^k : t_{(1:\|vt_i^k\|)}^k$	a k-valence utility-technology
$\|vt_0^{(k,i)}\|$	the valence of vt_i^k
$VT^k = \{vt_i^k \| 0 \leq i < \|VT^k\| - 1\}$	a set of k-valence utility-technologies
$HVT = \{VT^{(k)} \| 0 \leq k < \|HVT\| - 1\}$	hypothesis of valence theory
$\phi_i^k = vt_{(i,1:\|\phi_i^{(k)}\|-1)}^k \rightarrow vt_{(i,0)}^k$	a k-degree function-effect
$\|\phi_i^k\|$	the degree of ϕ_i^k
$\Phi^k = \{\phi_i^k \| 0 \leq i < \|\Phi^k\| - 1\}$	a set of k-degree function-effects
$FR = \{\Phi^k \| 0 \leq k < \|MVFR\| - 1\}$	function-effect representation

IV. Function-effect mining method

The function-effect representation(FR) is mined by the steps as shown in Algo.1, including 1) generation process of utility-technologies for obtaining functions/effcts in Algo.2, 2) generation process of transaction of functions/effcts in algo.3 and 3) mining association relation between functions and effects for obtaining function-effects of the patent by Eq.7.

In Algo.2, the valence relations between verb terms and noun terms are obtained by pos tagging and dependency parsing[2] from the claims and the abstract of a patent, which will generates different utility-technologies as functions/effects.

Given the functions/effcts, Algo.3 generates some transactions of the functions/effcts, where each sentence can be regards as a transaction consist of functions/effcts.

Given the transactions of functions/effcts, Algo.1 mines the relations of functions and effects with support and confidence large than some threshold values by Eq.7, resulting in the function-effect representation.

$$\left\{ vt_{(i,1:k-1)}^k \rightarrow vt_{i,0}^k \;\middle|\; \begin{array}{l} vt_{(i,1:k-1)}^k \bigcap vt_{(i,0)}^k = \emptyset \\ sup(vt_{(i,1:k-1)}^k \rightarrow vt_{(i,0)}^k) > \theta_s \\ conf(vt_{(i,1:k-1)}^k \rightarrow vt_{(i,0)}^k) > \theta_c \end{array} \right\}. \tag{7}$$

V. Experiments

In this section, we conduct some experiments to validate the effectiveness of our representation model.

[2]https://nlp.stanford.edu/software/

Algorithm 1: Mining Process of FR

Input: the abstract p^A, the claim p^C, the description p^D of a patent p
Output: the function-effects $FR(p)$ of p

initialize $FR = \emptyset$;
generate utility-technologies as functions/effects, HVT=Algo.2(p^A, p^C) ;
generate transactions of functions/effects, $Trans$=Algo.3(HVT, p^D) ;
if $\phi_i^k = vt^k_{(i,1:k-1)} \rightarrow vt^k_{(i,0)}$ *is consistent with Eq.7* **then**
  $\Phi^k = \Phi^k \bigcup \phi_i^k$;
end
return $FR = \{\Phi^k | 0 \le k < |FR| {-} 1\}$;

Algorithm 2: generation process of utility-technologies as functions/effects

Input: $p^A = \{s_i | 0 \le i < |p_A| {-} 1\}$,
$p^C = \{s_i | 0 \le i < |p_C| {-} 1\}$
Output: $HVT(p^A, p^C) = \{VT^k | 0 \le k < |HVT| {-} 1\}$

initialize $\{VT^{(k)} = \emptyset | 0 \le k < |HVT| {-} 1\}$;
for $s \in p^A \cup p^C$ **do**
  parse dependency tree $tree(s)$;
  if $\{t_{1:k}\}$ *directly dependent the same* v_0 *in* $tree(s)$ **then**
    $VT^k = VT^k \bigcup v_0 : t_{1:k}$;
  end
end
return utility-technologies $HVT = \{VT^k | 0 \le k < |HVT| {-} 1\}$;

Algorithm 3: generation process of function/effect transactions

Input: HVT, $p^D = \{s_i | 0 \le i < |p_D| {-} 1\}$
Output: $Trans = \{ts^{(k)} | 0 \le k < |Trans| {-} 1\}$

initialize transaction $Trans = \emptyset$;
for $s \in p^D$ **do**
  $ts = \emptyset$;
  for $vt \in VT$ *of* HVT **do**
    if $vt \subseteq s$ **then**
      $ts = ts \bigcup vt$;
    end
  end
  $Trans = Trans \bigcup ts$;
end
return $Trans$;

A. Experimental Datasets

Patent data are downloaded from U.S. Patent and Trademark Office(USPTO)[3], which is used in our experiments.

The patents used in our experiments are shown in table IV. The patent CPC codes of each patent are regarded its the multi-label, which are shown in table V. There are 4446 patents from 9 CPC codes.

TABLE IV
THE DESCRIPTION OF EXPERIMENTAL DATA

Source	USTPO								
Data Set	A	B	C	D	E	F	G	H	Y
#code	0	1	2	3	4	5	6	7	8
Number	1256	1508	504	65	279	746	2053	1594	211
Total Number	4446								

TABLE V
PATENT DATA WITH MULTI-LABEL

the number of patent	Multi-Label (ABCDEFGHY)
592	000000100
522	000000110
436	100000000
…	…

B. Baseline Models

We compare function-effect representation(FR) with following state-of-the-art representation models:

1) Vector Space Model (VSM)[1]: VSM is a concept-based model. For VSM, each patent is represented a vector, in which the word is encoded by one-hot.
2) Power Series Representation(PSR)[14]: PSR is a relation-based model. For PSR, each patent is represented by some association relations between technical terms.

C. Evaluation Measurements

Effective representation should have better performance in semantic clustering. The patents with the same CPC code are likely to cluster together. The clustering results are compared with the multi-labels of patent. We use a widely used evaluation measurements in our experiments. The precision, recall, F-measure are used to measure the class code predicted by the model with the reference codes.

Precision is calculated by,

$$P = \frac{TP}{TP + FP} \tag{8}$$

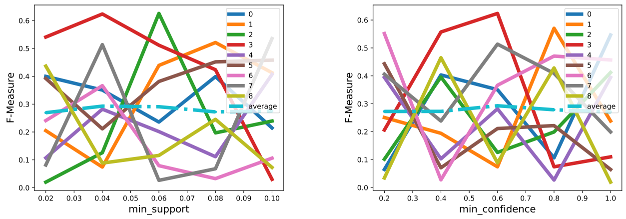
Recall is calculated by,

$$R = \frac{TP}{TP + FN} \tag{9}$$

F-measure is calculated by,

$$F = \frac{2 \times P \times R}{P + R} \tag{10}$$

[3]www.uspto.gov



(a) F-measure (b) F-measure

Fig. 2. (a) F-Measure under different support[0.02, 0.10], confidence 0.6. (b) F-Measure under support 0.04, different confidence[0.2, 1.0].

where TP denotes true positive; TN denotes true negative; FP denotes false positive; FN denotes false negative.

D. Experimental Setups

For each patent p in the data set, we make experiment as follows.

1) Represent the patent by function-effect representation(FR), vector space model (VSM) and Power Series Representation(PSR) in sectionV-B;
2) Cluster each patent by K-means method and each patent p are clustered into top k multi-clusters with most similarity(k equals to the number of different CPC codes in the patent multi-label);
3) Compare and evaluate the results on the evaluation measurement in the section V-C.

E. Experimental Results

The F-measure on 9 different CPC number codes under different support are shown in TableVI. The results are shown respectively in Fig.2 (a). For the experimental data, the highest average F-Measure of the 9 CPC number codes is obtained when the support is 0.04.

TABLE VI
F-MEASURE UNDER DIFFERENT SUPPORT IN RANGE[0.02,0.10] AND CONFIDENCE 0.6

Meas.	sup	0	1	2	3	4	5	6	7	8
F-Measure	0.02	0.3997	0.2048	0.0200	0.5408	0.1071	0.3921	0.2410	0.0815	0.4365
	0.04	0.3506	0.0742	0.1256	0.6237	0.2820	0.2113	0.3663	0.5137	0.0884
	0.06	0.2357	0.4390	0.6256	0.5115	0.2015	0.3814	0.0791	0.0265	0.1158
	0.08	0.3987	0.5211	0.1969	0.4235	0.1113	0.4517	0.0326	0.0680	0.2454
	0.10	0.2145	0.4128	0.2394	0.0300	0.4068	0.4580	0.1061	0.5355	0.0728

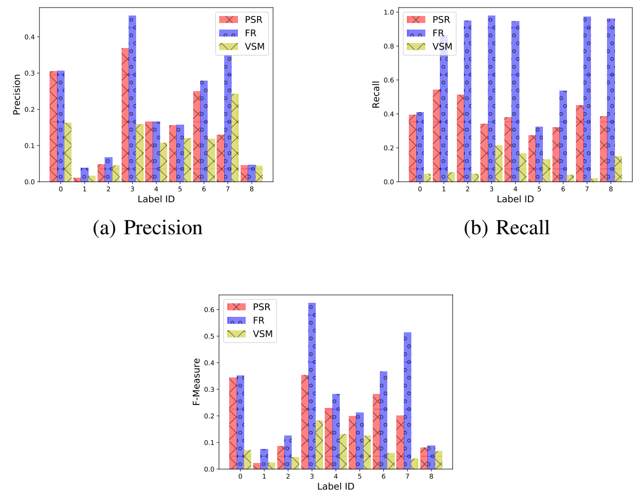
The F-Measure on 9 CPC number codes under different confidence are shown int TableVII. The results are shown respectively in Fig.2(b). For the experimental data, the highest average F-Measure of the 9 CPC number codes is obtained when the confidence is 0.6.

TABLE VII
F-MEASURE UNDER SUPPORT 0.04, DIFFERENT CONFIDENCE[0.2, 1.0]

Meas.	sup	0	1	2	3	4	5	6	7	8
F-Measure	0.2	0.0647	0.2508	0.1015	0.2061	0.3945	0.4433	0.5514	0.4061	0.0342
	0.4	0.4033	0.1942	0.3960	0.5572	0.1028	0.0707	0.0275	0.2381	0.4650
	0.6	0.3506	0.0741	0.1256	0.6237	2820	0.2113	0.3663	0.5136	0.0884
	0.8	0.1062	0.5705	0.1989	0.0742	0.0268	0.2218	0.4709	0.4089	0.4284
	1.0	0.5465	0.2383	0.4121	0.1098	0.3943	0.0646	0.4569	0.1991	0.0193

TABLE VIII
PRECISION, RECALL AND F-MEASURE FOR THREE MODELS

Meas.	Model	0	1	2	3	4	5	6	7	8
Precision	FR	0.3067	0.0387	0.0673	0.4578	0.1657	0.1566	0.2787	0.3491	0.0463
	PSR	0.3043	0.0111	0.0474	0.3681	0.1648	0.1555	0.2499	0.1290	0.0451
	VSM	0.1626	0.0156	0.0450	0.1580	0.1072	0.1202	0.1188	0.2426	0.0431
Recall	FR	0.4091	0.8642	0.9492	0.9778	0.9453	0.3249	0.5344	0.9716	0.9618
	PSR	0.3931	0.5420	0.5116	0.3401	0.3481	0.2727	0.3214	0.4507	0.3862
	VSM	0.0453	0.0544	0.0457	0.2133	0.1669	0.1303	0.0393	0.0207	0.1503
F-Measure	FR	0.3506	0.0742	0.1256	0.6237	0.2820	0.2113	0.3663	0.5137	0.0884
	PSR	0.3430	0.0218	0.0867	0.3536	0.2296	0.980	0.2812	0.2006	0.0808
	VSM	0.0709	0.0243	0.0453	0.1815	0.1305	0.1250	0.0590	0.0381	0.0670



(a) Precision (b) Recall

(c) F-Measure

Fig. 3. (a), (b), (c) are the Precision, Recall and F-Measure under different models.

When keep support 0.04 and confidence 0.6, the precision, recall and F-measure on the results obtained by our model compared with that of the baseline models. The comparative results are shown in table VIII.

Fig.3(a), (b) and (c) show the comparative results. The results show that our model outperforms the baseline models in precision, recall and F-measure.

VI. CONCLUSION

In this paper, we propose a function-effect based patent representation model. The contributions of our model are summarized as follow.

1) To represent the functions/effects in a patent, the function-effect representation which is inspired by hypothesis of valence theory, represents functions/effects by multi-valence utility-technologies;
2) To represent the function-effects in a patent, the function-effect representation, which is inspired by hypothesis of linearity concept, represents patent function-effects by multi-degree association relations between functions and effects.

Compared with the baseline models, the function-effect representation exhibits good performance in precise, recall and F-measures in the clustering task.

ACKNOWLEDGEMENT

The research work in this paper was supported by the National Science Foundation of China (grant no. 61801251) and Natural Science Foundation of Inner Mongolia (2018B-S06002).

REFERENCES

[1] G. Salton, A. Wong, and C.-S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[2] P. D. Turney and P. Pantel, "From frequency to meaning: Vector space models of semantics," *Journal of artificial intelligence research*, vol. 37, pp. 141–188, 2010.

[3] E. Gabrilovich, S. Markovitch *et al.*, "Computing semantic relatedness using wikipedia-based explicit semantic analysis." in *IJcAI*, vol. 7, 2007, pp. 1606–1611.

[4] G. Dobbie, W. Xiaoying, T. W. Ling, and M. L. Lee, "Ora-ss: An object-relationship-attribute model for semi-structured data," Tech. Rep., 2000.

[5] S. T. Dumais, "Latent semantic analysis," *Annual review of information science and technology*, vol. 38, no. 1, pp. 188–230, 2004.

[6] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin, "A neural probabilistic language model," *Journal of machine learning research*, vol. 3, no. Feb, pp. 1137–1155, 2003.

[7] A. Mnih and Y. W. Teh, "A fast and simple algorithm for training neural probabilistic language models," *arXiv preprint arXiv:1206.6426*, 2012.

[8] L. Azzopardi, M. Girolami, and M. Crowe, "Probabilistic hyperspace analogue to language," in *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, 2005, pp. 575–576.

[9] D. K. Tayal, L. Ahuja, and S. Chhabra, "Word sense disambiguation in hindi language using hyperspace analogue to language and fuzzy c-means clustering," in *Proceedings of the 12th International Conference on Natural Language Processing*, 2015, pp. 49–58.

[10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[11] X. Rong, "word2vec parameter learning explained," *arXiv preprint arXiv:1411.2738*, 2014.

[12] H. Nanba, H. Kamaya, T. Takezawa, M. Okumura, A. Shinmori, and H. Tanigawa, "Automatic translation of scholarly terms into patent terms," in *Proceedings of the 2nd international workshop on Patent information retrieval*, 2009, pp. 21–24.

[13] M. Okamoto, Z. Shan, and R. Orihara, "Applying information extraction for patent structure analysis," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2017, pp. 989–992.

[14] X. Luo, J. Zhang, F. Ye, P. Wang, and C. Cai, "Power series representation model of text knowledge based on human concept learning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 1, pp. 86–102, 2013.

[15] H. Zhuge, "Autonomous semantic link networking model for the knowledge grid," *Concurrency and Computation: Practice and Experience*, vol. 19, no. 7, pp. 1065–1085, 2007.

[16] X. Luo, Z. Xu, J. Yu, and X. Chen, "Building association link network for semantic link on web resources," *IEEE transactions on automation science and engineering*, vol. 8, no. 3, pp. 482–494, 2011.

[17] K. Nakayama, T. Hara, and S. Nishio, "Wikipedia link structure and text mining for semantic relation extraction." in *SemSearch*, 2008, pp. 59–73.

[18] K. Nakayama, M. Pei, M. Erdmann, M. Ito, M. Shirakawa, T. Hara, and S. Nishio, "Wikipedia mining," *Wikimania. Wikimedia*, 2008.

[19] K. Nakayama, T. Hara, and S. Nishio, "Wikipedia link structure and text mining for semantic relation extraction." in *SemSearch*, 2008, pp. 59–73.

[20] L. Zhang, L. Li, and T. Li, "Patent mining: a survey," *ACM Sigkdd Explorations Newsletter*, vol. 16, no. 2, pp. 1–19, 2015.

[21] S. Rogers, S. Tang, and J. Canny, "Acce: automatic coding composition evaluator," in *Proceedings of the first ACM conference on Learning@ scale conference*, 2014, pp. 191–192.

[22] G. A. Miller, "Wordnet: a lexical database for english," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[23] Z. Dong, Q. Dong, and C. Hao, "Hownet and the computation of meaning," 2006.

[24] Z. Wang, K. Zhao, H. Wang, X. Meng, and J.-R. Wen, "Query understanding through knowledge-based conceptualization," in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.