

Facilitating Technology Transfer by Patent Knowledge Graph

Weiwei Deng
 Department of Information
 Systems,
 City University of Hong Kong
wwdeng3-c@my.cityu.edu.hk

Xiaoming Huang
 Department of Information
 Systems,
 City University of Hong Kong
xmhuang6-c@my.cityu.edu.hk

Peihu Zhu
 Department of Information
 Systems,
 City University of Hong Kong
peihuzhu2-c@my.cityu.edu.hk

Abstract

Technologies are one of the most important driving forces of our societal development and realizing the value of technologies heavily depends on the transfer of technologies. Given the importance of technologies and technology transfer, an increasingly large amount of money has been invested to encourage technological innovation and technology transfer worldwide. However, while numerous innovative technologies are invented, most of them remain latent and un-transferred. The comprehension of technical documents and the identification of appropriate technologies for given needs are challenging problems in technology transfer due to information asymmetry and information overload problems. There is a lack of common knowledge base that can reveal the technical details of technical documents and assist with the identification of suitable technologies. To bridge this gap, this research proposes to construct knowledge graph for facilitating technology transfer. A case study is conducted to show the construction of a patent knowledge graph and to illustrate its benefit to finding relevant patents, the most common and important form of technologies.

1. Introduction

Technologies play an important role in driving the global development and the improvement of human well-being [3]. Lots of money is allocated to encourage innovation worldwide. A report in Nature shows that many countries invest more than 2% of their gross domestic product on research and development and the number is still growing [43]. Consequently, numerous technologies have been invented. Take patent, the most common and important form of technologies, as an example, the number of filed patent applications is increasingly large and reaches three million worldwide in 2016 as recorded by the World Intellectual Property Organization [47]. The value of technologies is realized through applications in practice. But practitioners

themselves usually do not have enough time and/or capital to develop innovative technologies. It is thus critical to transfer available technologies to the demand side. However, while innovative technologies are invented, most of them remain unused and many demands of practitioners stay unsatisfied. For instance, according to the inspection of National Congress of Korea in 2012, about 73% of the patents owned by universities and public-funded research institutions did not create social values [30]. Given this situation, it is essential to smooth the way of technology transfer.

Technology transfer offices in research institutions or companies have long been set up for facilitating technology transfer. But these divisions mainly depend on human labor. With the increasingly large number of technologies, it becomes difficult even impossible for humans to comprehend all technologies and identify suitable ones for a given demand [46]. Previous studies have explored various impacts [4, 12, 35] of and influential factors [13, 17, 20] in technology transfer. Some other studies have proposed to facilitate technology transfer from several different perspectives, such as finding experts for university-industry collaboration [45], building recommender systems for selective dissemination of research resources [33], and conducting various patent analysis for understanding technology development and technology transfer [29, 39, 49]. However, the information asymmetry between technology inventors and adaptors and the information overload problem remain challenging to technology transfer. The demand side lacks detailed understanding of technical documents, and the large number of technical documents further impedes the identification of suitable technologies for a given need. This research attempts to facilitate technology transfer from the perspective of better comprehension and more accurate identification of technologies with the support of knowledge graph. Knowledge graph is a graph structured knowledge base that stores factual information in form of semantic relations between entities [27]. It can automatically extract core entities and their semantic relationships from documents and represents them in a machine-understandable way, and

therefore is helpful for overcoming the information asymmetric and information overload problems.

The main objective of this research is to construct a patent knowledge graph for facilitating technology transfer. Natural language processing techniques and knowledge extraction techniques can be employed to process patent documents and construct the patent knowledge graph automatically. Specifically, a graph schema is first defined to clarify the types of entities and relationships that are included in the patent knowledge graph. A small sample set of patent documents is annotated manually with possible entities and relationships among them. The annotated documents are fed to a conditional random field model for training. The trained model is then employed to automatically extract entity instances and their relationships from other patent documents. The extracted entity instances and relationships are finally stored in the format of subject-predicate-object triples. By constructing the patent knowledge graph, the main information of patent documents becomes explicit and machine-readable. Consequently, the constructed patent knowledge graph can not only facilitate the comprehension of complicated technical documents, but also assist with the identification of needed technology. A case study in digital data processing domain is conducted and used to demonstrate the advantages of patent knowledge graph in facilitating technology transfer. Overall, this research advances the technology transfer literature by proposing a knowledge graph approach and contributes to the patent analysis literature by using knowledge graph to extract and represent the detailed and semantic information of patent documents.

The remaining of this paper is organized as follows. The second section reviews studies related to technology transfer and knowledge graph, and identifies the research gap. The third section provides the procedure of constructing a patent knowledge graph. The next section presents a case study in digital data processing area and illustrates the benefit of patent knowledge graph to technology transfer. And the last section concludes this research with its main contributions and implications.

2. Related work

2.1. Technology transfer

Technology transfer is a long studied area in both academia and industry. In this research, technology transfer refers to the movement of know-how, technical knowledge, or technology from one organizational setting to another [9]. One of the main streams of technology transfer research is to identify and explore

factors that can accelerate or hamper the process. During the process, four factors are identified and considered to play important roles. First of all, communication interactivity, which refers to interactions between technology inventors and receptors, can improve the probability of successful technology transfer. Second, cultural and geographical distances can inhibit technology transfer as they bring more difficulties to technology inventors and receptors for achieving an agreement on patents' value, practices and so on. In addition, as stated by Albrecht and Ropp [1], cultural difference is viewed more important on making sense of distance. Third, technology equivocality, which is defined as the level of concreteness of technology [9, 10], has a negative effect on technology transfer. That is because technology with high equivocality is more difficult for adopters to understand and to put into practice. The last one is personal motivation: the greater motivation of technology inventors to diffuse and of technology receivers to adopt makes technology transfer more likely to occur [9]. Reward is one of the most common methods to improve the personal incentives, such as tax free policy and rebate. To sum up, three of the four factors (communication interactivity, distance, and technology equivocality) are, to certain extent, related to the understanding of technologies. In addition, the increasingly large number of technologies also brings great difficulty to effective technology transfer [33].

Another research stream related to technology transfer is patent analysis. Many researchers have conducted various patent analysis, such as patent evaluation, technology trend analysis, and patent classification. Patent evaluation facilitates technology transfer by identifying potential high-quality patents. And many studies are aimed at identifying quality indicators. For instance, Gerken and Moehle [8] evaluated the level of novelty of patents through semantic analysis. Trappey et al. [41] used International Patent Classification (IPC) and the number of citations as indicators to evaluate patent quality. A greater number of forward citations suggests a bigger commercial interest, and a larger number of backward citations indicates higher validity of patents [42]. In addition, the number of IPC classes is argued to represent the broadness of a patent. More classes indicates greater scope and thus higher value [21, 26]. Technology trend analysis helps to identify promising patents that are more likely to be adopted. Hence, it attracts much attention from researchers. For example, Kim et al. [14] identified emerging technology by building a semantic network of keywords considering both structured and unstructured content. Park and his colleagues [28] identified promising patents through analyzing TRIZ evolution trends. Yoon and Kim [49]

identified technology trends through extracting and analyzing the properties and functions of technologies. Patent classification also contributes to technology transfer in certain extent. Classifying patents into pre-designed categories makes patent searching and retrieving much easier, which thus improves the success of technology transfer. Liu et al. [23] developed the Patent Retrieval and Analysis Platform, which matches fields of patents and identifies similar patents based on bibliographic pattern discovery and text mining approaches. Shih and Liu [38] proposed a network based classification approach that firstly constructs ontology network and then identifies k-nearest neighbors and patent classes. Patents' citations and content features are also frequently used to classify patents [22, 39].

For technology transfer, it is important to understand the content of technical documents as well as identifying suitable technologies when needed. Although much attention has been paid to patent evaluation, technology trend analysis, and patent classification, little effort has been made to comprehend the technical details of technical documents and leverage the technical details to identify needed technologies. Therefore, in this research, we attempt to develop a patent knowledge graph that can reveal the semantic knowledge embedded in patent documents and assist with the identification of suitable patents.

2.2. Knowledge graph

The concept of knowledge graph has gained much attention since Google launched its knowledge graph in 2012 [2]. A knowledge graph is a graph structured knowledge base in which knowledge is represented by relationships between entities [27, 31, 48]. Due to the properties of effective information integration, machine-readable knowledge, and comprehensive entity summarization, knowledge graph has been widely used to support knowledge-intensive applications, such as information retrieval [2], automatic question answering [24], personalized recommendation [6], and technology trend prediction [7]. Different applications require different specialized knowledge. Therefore, knowledge graphs are usually customized for specific applications.

The construction of knowledge graph requires two critical components: knowledge graph schema and knowledge extraction techniques. Schema specifies the types of entities and relationships to be included in the knowledge graph. It is defined according to a specific application. Given a schema, knowledge extraction techniques are used to extract entity instances and their relationships from various data sources. In terms of textual data, there are two major types of knowledge extraction techniques, namely rule-based techniques

and statistical learning-based techniques [36]. Rule-based techniques extract knowledge based on predefined rules and perform well when the target knowledge is related to certain language patterns [15]. However, defining rules requires large expert efforts and the rules may not be generalizable to a larger set of textual data [34]. On the other hand, statistical learning-based techniques extract knowledge based on mathematical models learned from training samples. They use statistical methods to produce their own rules and classifiers that are more generalizable. But a large training set is needed for producing a satisfactory model [11]. Between these two types of techniques, statistical learning-based methods are more commonly used since constructing knowledge graph usually involves complex and large volume of textual data. Among various statistical learning-based techniques, conditional random fields (CRFs) are of wide popularity due to the strength of incorporating rich and overlapping layout and language features. Given these advantages, CRFs have been widely applied to extract knowledge from textual data [5, 19, 32, 37]. In this study, therefore, we employ the CRF model proposed by Lee et al [19] to extract knowledge from patent documents.

3. Patent knowledge graph construction

The procedure of constructing a patent knowledge graph comprises three major steps. The first step is to define the patent knowledge graph schema that specifies entities and relationships important to understanding technical documents. The second step is to create a set of labeled data by manually annotating patent documents with entity types and relationships. This step is not necessary if there is already enough labeled data. The last step is to train knowledge extraction model with the labeled data and use the trained model to extract possible entity instances and relationships from new patent documents. The extracted entities and relationships are then stored in the format of subject-predicate-object triples which constitute the patent knowledge graph. The details of these steps are introduced in the following subsections.

3.1. Patent knowledge graph schema

With the aim of facilitating technology transfer, we define the patent knowledge graph schema as follows:

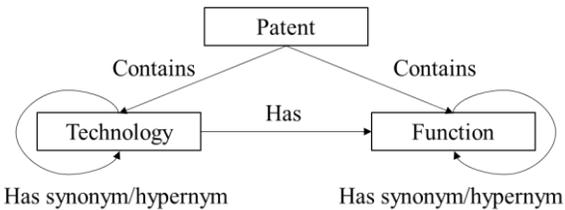


Figure 1. The patent knowledge graph schema

As indicated by the theory of inventive problem solving [40, 44], technology and function are two essential components in patent documents. A patent contains one or more technologies which could be new solutions, systems, products, etc. Each technology has one or more functions that could be its attributes, objectives, capabilities, and so on. It is well acknowledged that comprehending patent documents is difficult because of the complex linguistic style [46]. To facilitate the comprehension of patents, therefore, we further consider the synonyms and hypernyms of technologies and functions. In brief, given a patent document, we mainly extract its technologies, functions, the dependency relationship between functions and technologies, and their synonym or hypernym relationships. There are three types of patents, namely utility patents, design patents, and plant patents. However, design and plant patents do not contain the technical aspects defined in the proposed schema. The proposed schema is biased towards utility patents since 90 percent of all patents are utility patents according to the USPTO (United States Patent and Trademark Office).

3.2. Annotation

Training samples are needed for automatic knowledge extraction because the CRF model used in this research is a supervised technique. Since there is no labeled data available in the patent domain, manual annotation is needed to create a training set based on the defined patent knowledge graph schema. Specifically, given a piece of textual data, words or phrases that belong to certain entity types are annotated with corresponding entity types. Relationships between the identified words or phrases are also annotated if there is any. Figure 2 presents an example of annotated textual data.

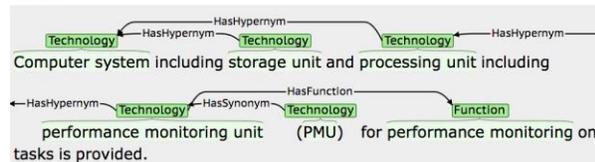


Figure 2. An annotation example

3.3. Knowledge extraction

Knowledge extraction has two key parts: entity recognition and relationship extraction. Entity recognition tries to decide whether words or phrases belong to certain types of entities. Relationship extraction deals with the task of detecting and classifying relationships between identified entities. These two tasks can be performed simultaneously using CRFs. CRFs are undirected graphical models used to encode known relationships between observations and build interpretation models [16]. The CRF model employed in this research is a linear-chain CRF which is effective in predicting the sequence of labels for an input sequence [19].

Formally, the CRF model is defined as follows. Given a piece of textual data, let $o = \{o_1, o_2, \dots, o_n\}$ be the sequence of observed words in the textual data, and $s = \{s_1, s_2, \dots, s_n\}$ be the sequence of states which correspond to labels assigned to each word in o . Given an input sequence o , the conditional probability of state sequence s is defined below:

$$P(s|o) = \frac{1}{Z_o} \exp\left(\sum_{i=1}^n \sum_{j=1}^m \lambda_j f_j(s_{i-1}, s_i, o, i)\right), \quad (1)$$

$$Z_o = \sum_{s \in S^n} \exp\left(\sum_{i=1}^n \sum_{j=1}^m \lambda_j f_j(s_{i-1}, s_i, o, i)\right), \quad (2)$$

where Z_o is a normalization factor used to ensure the sum of the probabilities of all state sequences to be one, S^n is the set of all possible state sequences, m is the number of features, n is the length of input sequence, $f_j(s_{i-1}, s_i, o, i)$ is a binary feature function, and λ_j is a learned weight indicating the preference on feature j . The parameters of the CRF model are estimated by maximizing the conditional probability of a set of labeled samples. Given a new and unlabeled input sequence, the most possible state sequence is then returned based on the trained model.

In the CRF model, the following features are considered with the support of Stanford CoreNLP toolkit [25]: word (the original form of input words), lemma (the lemma of a word with respect to its intended meaning), part-of-speech (such as noun, verb, and adjective), and syntactic phrase (one type of syntactic unit in the grammar structure, such as noun phrases).

4. A case study

A case study is conducted to demonstrate the effectiveness of patent knowledge graph in facilitating technology transfer. Specifically, we first construct a patent knowledge graph related to the domain of digital data processing. The constructed patent knowledge graph is then applied to facilitate the identification of suitable technologies for given needs. Details of this case study are introduced in the following subsections.

4.1. Constructing patent knowledge graph

We randomly select 5000 US (United States) patents granted in 2017 from PatentsView database¹. All selected patents are related to ‘Electric Digital Data Processing’, the ‘G06F’ subclass of the International Patent Classification (IPC). When extracting knowledge from patents, only titles and abstract of patents are considered because they summarize the main content of patents.

Among the 5000 selected patents, we further select 300 patents randomly for annotation. The annotated patents are then used as training set to train the CRF model. The trained CRF model is finally used to extract possible technologies, functions, and relationships among them from all the selected patents. In total, we extract 5050 unique technology instances, 6981 unique function instances, 12570 unique ‘has function’ relationship instances, 332 unique hypernym relationship instances, and 9 unique synonym relationship instances. The unique instances of hypernym relationship and synonym relationship are relatively few especially for the latter. This is because there are only a few such relationship instances in patent titles and abstracts. Consequently, few such instances are annotated in the training set. With a very small number of annotated instances, the CRF model is unlikely to be trained well for extracting such instances from new documents.



Figure 3. Part of the patent knowledge graph

Since the whole patent knowledge graph is too dense to present in one static figure, we present part of the constructed patent knowledge graph in Figure 3. This

¹ <http://www.patentsview.org/download/>

figure contains 20 randomly selected patents and their related technologies and functions. We further present the details of the constructed patent knowledge graph using three examples. Figure 4 presents a sub-graph of a patent whose patent number is ‘9852480’. It is easy to understand from this sub-graph that this patent contains two technologies (i.e., data processing system and activity management system) and four functions (i.e., managing activities linked to multimedia content, provide a platform, present viewers with multimedia content, and access multimedia content). Such knowledge can facilitate our comprehension of complicated technical documents.

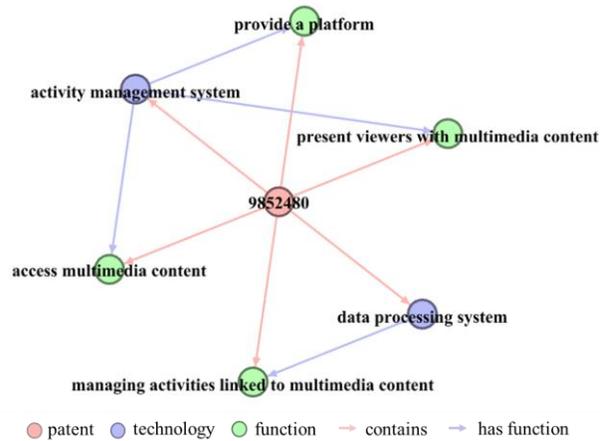


Figure 4. A sub-graph of a patent

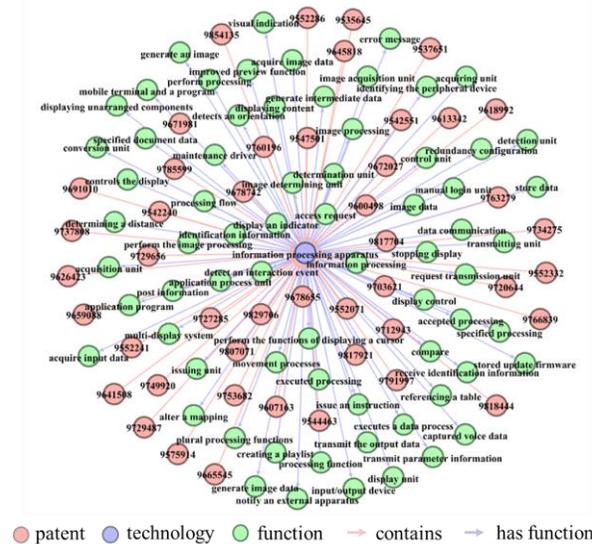


Figure 5. A sub-graph of a technology

We further present a sub-graph related to ‘information processing apparatus’ technology in Figure 5. Given a technology, it is very convenient to

summarize its related patents and functions. In this case, 46 patents are identified to contain the ‘information processing apparatus’ technology, and the technology has 69 related functions.

Finally, we present a sub-graph related to the ‘access request’ function in Figure 6. From the whole patent knowledge graph, we identify 11 patents that contain the ‘access request’ function and 17 technologies which contain that function. Both technology and function centered sub-graphs can help us identify relevant patents for given needs. Besides, the two kinds of graphs can also facilitate our understanding of given technologies or functions. In short, patent knowledge graph can provide rich semantic knowledge for better understanding of patent documents as well as identification of relevant patents.

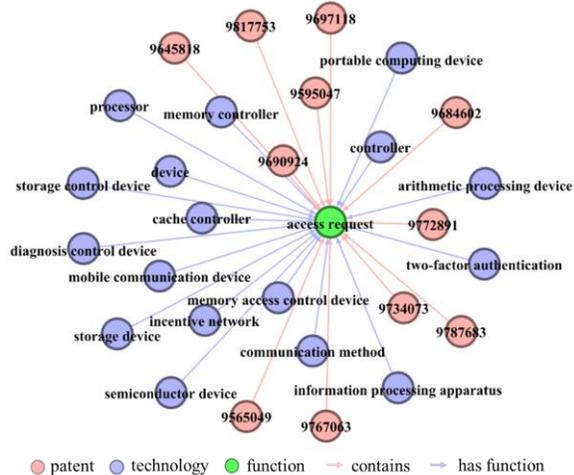


Figure 6. A sub-graph of a function

4.2. Application of the patent knowledge graph

To further illustrate the effectiveness of patent knowledge graph in identifying suitable technologies for given needs, we conduct an experiment on patent retrieval. The data set used in the patent retrieval experiment is also extracted from the PatensView database. Specifically, we randomly select 115 US patents which are granted in 2017 and belong to the ‘G06F’ subclass. We further extract patents that are cited by these 115 patents. Among all the cited patents, only US patents granted in 2017 and located in the ‘G06F’ subclass are remained. Consequently, 345 cited patents are identified. In summary, 610 citation records are extracted, each of the 115 patents has at least 3 cited patents. For patent retrieval experiment, the titles and abstracts of the 115 selected patents are considered as given needs, the patents cited by each of the selected patents are treated as the suitable technologies for corresponding needs. All patent titles and abstracts are processed by the Stanford CoreNLP toolkit so as to

remove stop words, standardize derivative words, and remain informative keywords.

A commonly used retrieval method is vector space model [18] which represents text documents and queries as keyword vectors:

$$d_j = (k_{j,1}, k_{j,2}, \dots, k_{j,t}), \quad (3)$$

$$q = (k_{q,1}, k_{q,2}, \dots, k_{q,t}), \quad (4)$$

where $k_{j,t}$ (respectively $k_{q,t}$) is a binary value that equals 1 if keyword t appears in document j (respectively query q) and 0 otherwise. Given a need and a pool of patents, relevant patents can be identified by calculating the cosine similarity between their keyword vectors using Equation 5. Patents with the highest similarities are then returned for the need.

$$sim(q, d_j) = \frac{\sum_{l=1}^t k_{q,l} \times k_{j,l}}{\sqrt{\sum_{l=1}^t k_{q,l}^2} \times \sqrt{\sum_{l=1}^t k_{j,l}^2}}. \quad (5)$$

The traditional retrieval method suffers from mismatch problem because semantic relations between keywords are ignored. In other words, patents that are semantically related to queries but do not contain the exact keywords are filtered out. The constructed patent knowledge graph contains rich semantic knowledge and thus can help identify relevant patents for a given need. To embed the semantic knowledge of patent knowledge graph into the retrieval process, we propose to expand query keywords with semantically related keywords from the patent knowledge graph. Specifically, given a need in its original form, possible technologies and functions are first identified by the trained CRF model. For each identified entity, its related concepts in one-hop range in the constructed patent knowledge graph are extracted and considered as the expanded need. To make sure that concepts extracted from the patent knowledge graph are relevant to the original need, we only consider entities which co-occur with the identified entity for at least certain number of times. For simplicity, in this experiment, the threshold is set to 2. The expanded need is processed following the same text processing procedure to obtain the processed keywords, which are then added to the keyword vector of the original need. The expanded keyword vector is represented as below:

$$q_{expanded} = (k_{q,1} + f_1, k_{q,2} + f_2, \dots, k_{q,t} + f_t), \quad (6)$$

where $k_{q,t}$ is the keyword value in the original query and f_t is the number of times that keyword t is added to the original query. The expanded keyword vector of the need is finally matched with keyword vectors of the pool of patents. Patents with the highest similarities are returned for the need.

To compare the retrieval performance with and without the patent knowledge graph, the following three measures are used:

$$precision = \frac{|TS \cap RS|}{|RS|}, \quad (7)$$

$$recall = \frac{|TS \cap RS|}{|RS|}, \quad (8)$$

$$F_measure = \frac{2 \times precision \times recall}{precision + recall}, \quad (9)$$

where TS is a test set that contains patents actually cited by the patent defining the given need, RS is a retrieved set which includes patents returned for the given need, and $|RS|$ (or $|TS|$) indicates the number of patents in the retrieved set (or test set).

Figure 7, Figure 8, and Figure 9 presents the patent retrieval performance in terms of precision, recall, and F-measure respectively. In these figures, traditional method stands for the retrieval method without the support of the patent knowledge graph. KG-based method refers to the retrieval method supported by the patent knowledge graph. These results show that the knowledge graph-based method outperforms the traditional retrieval method.

Table 1 shows the improvement of the knowledge graph-based method compared to the traditional method. As shown in the table, the improvement is relatively small when only a small number (1, 2, or 3) of patents are returned. But when more patents are returned, the proposed method gains much higher improvement. Paired t-test is conducted to further evaluate whether the improvement is significant or not. The result shows that the inclusion of patent knowledge graph significantly improves the patent retrieval performance when the number of returned patents is more than 3. Traditional keyword matching method can retrieve a few relevant patents, but other relevant patents without the exact keywords are ignored. The patent knowledge graph provides semantic knowledge to the original query and therefore helps identify more patents relevant to the query.

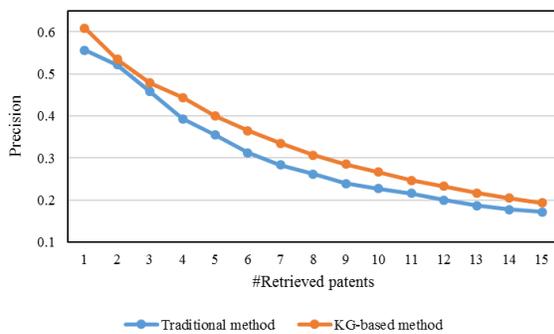


Figure 7. Precision of patent retrieval

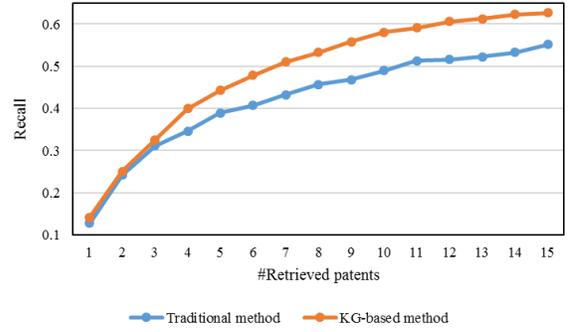


Figure 8. Recall of patent retrieval

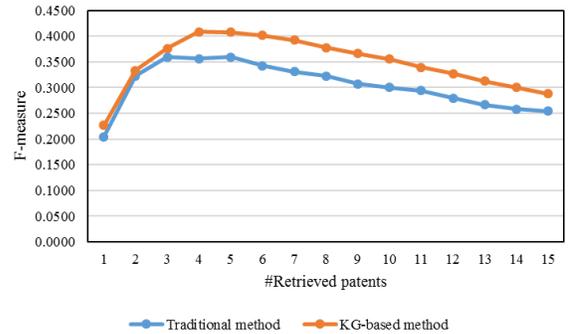


Figure 9. F-measure of patent retrieval

Table 1. Improvement of KG-based method

#Retrieved patents	Precision	Recall	F-measure
1	8.6%	9.8%	9.8%
2	2.4%	3.2%	3.1%
3	4.2%	4.5%	4.5%
4	11.3% *	13.3% **	12.7% **
5	11.3% **	12.0% **	11.9% **
6	14.3% ***	15.0% ***	14.9% ***
7	15.6% ***	15.2% ***	15.6% ***
8	14.5% ***	14.3% ***	14.6% ***
9	15.9% ***	16.0% ***	16.2% ***
10	15.0% ***	15.6% ***	15.6% ***
11	12.8% ***	13.3% ***	13.3% ***
12	14.0% ***	14.8% ***	14.6% ***
13	13.8% ***	14.7% ***	14.6% ***
14	13.4% ***	14.3% ***	14.1% ***
15	11.4% ***	11.9% ***	11.8% ***

Note: the asterisk label indicates that the improvement is significant in paired t-test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

5. Conclusion

This research attempts to facilitate technology transfer from the perspective of better comprehension of

technical documents and more accurate identification of technologies for given needs. To this aim, we propose to construct and apply a patent knowledge graph. Knowledge graph, on the one hand, can automatically extract semantic and machine-readable knowledge from technical documents. On the other hand, knowledge graph can provide rich knowledge for facilitating the understanding and identification of needed technologies. A case study is conducted to demonstrate the effectiveness of patent knowledge graph in facilitating technology transfer. We first construct a patent knowledge graph related to the digital data processing domain. The constructed patent knowledge graph is presented and discussed with several examples. We then employ the constructed patent knowledge graph in a patent retrieval task. A knowledge graph-based retrieval method is proposed to embed semantic information from the knowledge graph into the retrieval process. An experiment is conducted with real-world data and the result shows that the proposed method significantly outperforms the traditional retrieval method. Overall, the case study demonstrates the benefit of patent knowledge graph to technology transfer.

In addition, there are several parts that can be further researched in the future. The first one is to build a more comprehensive patent knowledge graph. Current study extracts technology, functions, and three kinds of relationships among them. Other information, such as patent citation and classification, may also be meaningful to technology transfer. Future work will survey other patent information that is important to technology transfer and accordingly construct a better patent knowledge graph. Second, there could be other ways to embed knowledge from patent knowledge graph into the patent retrieval process. We will investigate other knowledge embedding mechanisms and compare their performance in patent retrieval task. Third, the constructed patent knowledge graph can also be applied in other tasks, such as patent recommendation and technology trend analysis. These future studies can provide further support to the effectiveness of patent knowledge graph in facilitating technology transfer.

References

- [1] Albrecht, T.L., and V.A. Ropp, "Communicating about Innovation in Networks of Three U.S. Organizations", *Journal of Communication* 34(3), pp. 78–91.
- [2] Amit Singhal, "Introducing the Knowledge Graph: things, not strings", *Google*, 2012. <https://www.blog.google/products/search/introducing-knowledge-graph-things-not/>
- [3] Audretsch, D.B., E.E. Lehmann, and M. Wright, "Technology transfer in a global economy", *The Journal of Technology Transfer* 39(3), 2014, pp. 301–312.
- [4] Bozeman, B., H. Rimes, and J. Youtie, "The evolving state-of-the-art in technology transfer research: Revisiting the contingent effectiveness model", *Research Policy* 44(1), 2015, pp. 34–49.
- [5] Bundschuh, M., M. Dejori, M. Stetter, V. Tresp, and H.-P. Kriegel, "Extraction of semantic biomedical relations from text using conditional random fields", *BMC Bioinformatics* 9, 2008, pp. 207.
- [6] Catherine, R., K. Mazaitis, M. Eskenazi, and W. Cohen, "Explainable Entity-based Recommendations with Knowledge Graphs", *arXiv:1707.05254 [cs]*, 2017.
- [7] Duan, W., and Y.-Y. Chiang, "Building Knowledge Graph from Public Data for Predictive Analysis: A Case Study on Predicting Technology Future in Space and Time", *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*, ACM (2016), 7–13.
- [8] Gerken, J.M., and M.G. Moehrle, "A new instrument for technology monitoring: novelty in patents measured by semantic patent analysis", *Scientometrics* 91(3), 2012, pp. 645–670.
- [9] Gibson, D.V., and R.W. Smilor, "Key variables in technology transfer: A field-study based empirical analysis", *Journal of Engineering and Technology Management* 8(3), 1991, pp. 287–312.
- [10] Gibson, D.V., and R.W. Smilor, *Technology Transfer in Consortia and Strategic Alliances*, Rowman & Littlefield, 1992.
- [11] Guangpu, F., C. Xu, and P. Zhiyong, "A Rules and Statistical Learning Based Method for Chinese Patent Information Extraction", *2011 Eighth Web Information Systems and Applications Conference*, (2011), 114–118.
- [12] Han, J.-S., and S.-Y.T. Lee, "The impact of technology transfer contract on a firm's market value in Korea", *The Journal of Technology Transfer* 38(5), 2013, pp. 651–674.
- [13] Hsu, D.W.L., Y.-C. Shen, B.J.C. Yuan, and C.J. Chou, "Toward successful commercialization of university technology: Performance drivers of university technology transfer in Taiwan", *Technological Forecasting and Social Change* 92, 2015, pp. 25–39.
- [14] Kim, Y.G., J.H. Suh, and S.C. Park, "Visualization of patent analysis for emerging technology", *Expert Systems with Applications* 34(3), 2008, pp. 1804–1812.
- [15] Kluegl, P., M. Toepfer, P.-D. Beck, G. Fette, and F. Puppe, "UIMA Ruta: Rapid development of rule-based information extraction applications", *Natural Language Engineering* 22(1), 2016, pp. 1–40.
- [16] Lafferty, J., A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data", *Departmental Papers (CIS)*, 2001.
- [17] Lai, W.-H., "Willingness-to-engage in technology transfer in industry–university collaborations", *Journal of Business Research* 64(11), 2011, pp. 1218–1223.
- [18] Larson, R.R., "Introduction to Information Retrieval", *Journal of the American Society for Information Science and Technology* 61(4), pp. 852–853.
- [19] Lee, L.-H., K.-C. Lee, and Y.-H. Tseng, "The NTNU System at SemEval-2017 Task 10: Extracting Keyphrases and Relations from Scientific Publications Using Multiple Conditional Random Fields", *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-*

- 2017), Association for Computational Linguistics (2017), 951–955.
- [20] Leischnig, A., A. Geigenmueller, and S. Lohmann, “On the role of alliance management capability, organizational compatibility, and interaction quality in interorganizational technology transfer”, *Journal of Business Research* 67(6), 2014, pp. 1049–1057.
- [21] Lerner, J., “The Importance of Patent Scope: An Empirical Analysis”, *The RAND Journal of Economics* 25(2), 1994, pp. 319–333.
- [22] Li, X., H. Chen, Z. Zhang, J. Li, and J.F. Nunamaker, “Managing Knowledge in Light of Its Evolution Process: An Empirical Study on Citation Network-Based Patent Classification”, *Journal of Management Information Systems* 26(1), 2009, pp. 129–154.
- [23] Liu, S.-H., H.-L. Liao, S.-M. Pi, and J.-W. Hu, “Development of a Patent Retrieval and Analysis Platform – A hybrid approach”, *Expert Systems with Applications* 38(6), 2011, pp. 7864–7868.
- [24] Lukovnikov, D., A. Fischer, J. Lehmann, and S. Auer, “Neural Network-based Question Answering over Knowledge Graphs on Word and Character Level”, *Proceedings of the 26th International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee (2017), 1211–1220.
- [25] Manning, C., M. Surdeanu, J. Bauer, J. Finkel, S. Bethard, and D. McClosky, “The Stanford CoreNLP Natural Language Processing Toolkit”, *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Association for Computational Linguistics (2014), 55–60.
- [26] Merges, R.P., and R.R. Nelson, “On the Complex Economics of Patent Scope”, *Columbia Law Review* 90(4), 1990, pp. 839–916.
- [27] Nickel, M., K. Murphy, V. Tresp, and E. Gabrilovich, “A Review of Relational Machine Learning for Knowledge Graphs”, *Proceedings of the IEEE* 104(1), 2016, pp. 11–33.
- [28] Park, H., J.J. Ree, and K. Kim, “Identification of promising patents for technology transfers using TRIZ evolution trends”, *Expert Systems with Applications* 40(2), 2013, pp. 736–743.
- [29] Park, H., J. Yoon, and K. Kim, “Using function-based patent analysis to identify potential application areas of technology for technology transfer”, *Expert Systems with Applications* 40(13), 2013, pp. 5260–5265.
- [30] Park, Y.-J., “Social Tagging-based Recommendation Platform for Patented Technology Transfer”, *Journal of Intelligence and Information Systems* 21(3), 2015, pp. 53–77.
- [31] Paulheim, H., “Knowledge graph refinement: A survey of approaches and evaluation methods”, *Semantic Web* 8(3), 2017, pp. 489–508.
- [32] Peng, F., and A. McCallum, “Information extraction from research papers using conditional random fields”, *Information Processing & Management* 42(4), 2006, pp. 963–979.
- [33] Porcel, C., A. Tejada-Lorente, M.A. Martínez, and E. Herrera-Viedma, “A hybrid recommender system for the selective dissemination of research resources in a Technology Transfer Office”, *Information Sciences* 184(1), 2012, pp. 1–19.
- [34] Reiss, F., S. Raghavan, R. Krishnamurthy, H. Zhu, and S. Vaithyanathan, “An Algebraic Approach to Rule-Based Information Extraction”, *2008 IEEE 24th International Conference on Data Engineering*, (2008), 933–942.
- [35] Roessner, D., J. Bond, S. Okubo, and M. Planting, “The economic impact of licensed commercialized inventions originating in university research”, *Research Policy* 42(1), 2013, pp. 23–34.
- [36] Sarawagi, S., “Information Extraction”, *Foundations and Trends® in Databases* 1(3), 2008, pp. 261–377.
- [37] Settles, B., “Biomedical Named Entity Recognition Using Conditional Random Fields and Rich Feature Sets”, *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications*, Association for Computational Linguistics (2004), 104–107.
- [38] Shih, M.-J., and D.-R. Liu, “Patent Classification Using Ontology-Based Patent Network Analysis”, *PACIS 2010 Proceedings*, 2010.
- [39] Taduri, S., G.T. Lau, K.H. Law, and J.P. Kesan, “A Patent System Ontology for Facilitating Retrieval of Patent Related Information”, *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, ACM (2012), 146–157.
- [40] Terninko, J., A. Zusman, and B. Zlotin, *Systematic Innovation: An Introduction to TRIZ (Theory of Inventive Problem Solving)*, CRC Press, 1998.
- [41] Trappey, A.J.C., C.V. Trappey, C.-Y. Wu, C.Y. Fan, and Y.-L. Lin, “Intelligent patent recommendation system for innovative design collaboration”, *Journal of Network and Computer Applications* 36(6), 2013, pp. 1441–1450.
- [42] Trappey, A.J.C., C.V. Trappey, C.-Y. Wu, and C.-W. Lin, “A patent quality analysis for innovative technology and product development”, *Advanced Engineering Informatics* 26(1), 2012, pp. 26–34.
- [43] Van Noorden, R., “China by the numbers”, *Nature News* 534(7608), 2016, pp. 452.
- [44] Vincent, J.F.V., and D.L. Mann, “Systematic technology transfer from biology to engineering”, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 360(1791), 2002, pp. 159–173.
- [45] Wang, Q., J. Ma, X. Liao, and W. Du, “A context-aware researcher recommendation system for university-industry collaboration on R&D projects”, *Decision Support Systems* 103, 2017, pp. 46–57.
- [46] Wanner, L., R. Baeza-Yates, S. Brüggemann, et al., “Towards content-oriented patent document processing”, *World Patent Information* 30(1), 2008, pp. 21–33.
- [47] World Intellectual Property Organization, *World intellectual property indicators*, WIPO, Geneva, 2017.
- [48] Xiong, C., R. Power, and J. Callan, “Explicit Semantic Ranking for Academic Search via Knowledge Graph Embedding”, *Proceedings of the 26th International Conference on World Wide Web*, International World Wide Web Conferences Steering Committee (2017), 1271–1279.
- [49] Yoon, J., and K. Kim, “TrendPerceptor: A property-function based technology intelligence system for identifying technology trends from patents”, *Expert Systems with Applications* 39(3), 2012, pp. 2927–2938.