An Approach for Identifying Complementary Patents Based on Deep Learning

Jinzhu Zhang, Jialu Shi

Jialu Shi

Master student

Nanjing University of Science and Technology

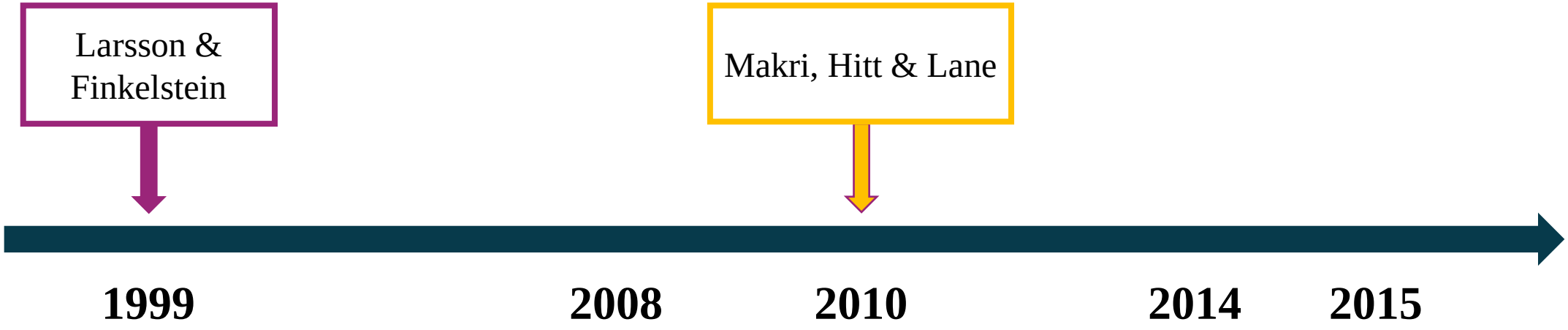
Contents

Introduction

Patent complementarity

Within the same broad scope of knowledge, the degree of difference in technical knowledge between the two technology R&D subjects in solving problems focused on different narrow scope of knowledge areas.

The degree to which two patent subjects with the same broad field of technology focus on different narrow domain technologies.



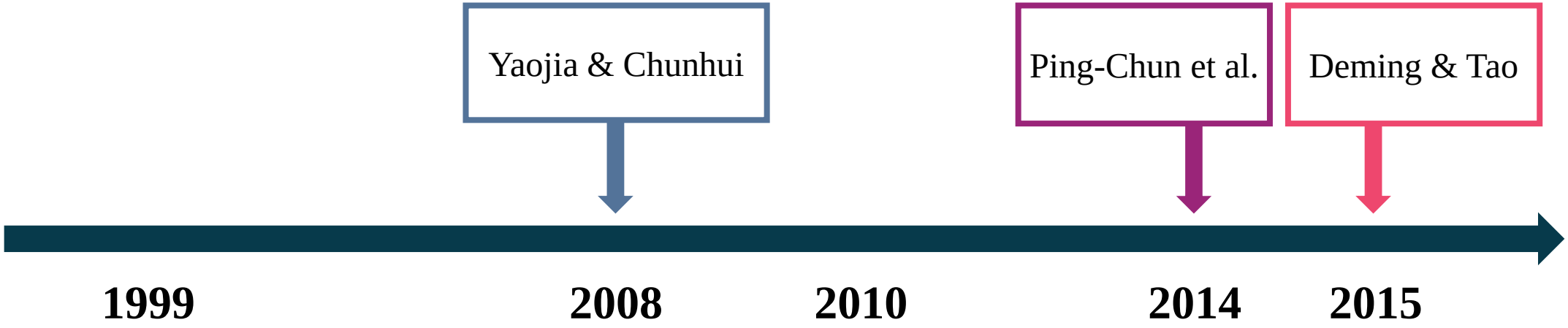
Introduction

Patent complementarity

Complementary relationship between two patents that are necessary and irreplaceable for the production of a product.

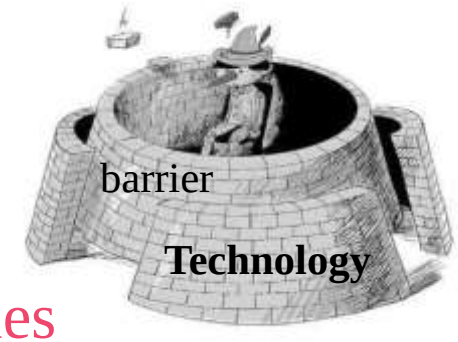
Patent citation represents direct dependence of technology, and complementarity is considered to exist when there is a double indirect citation relationship between patents and when the knowledge shared with each other is less similar.

Two patents are considered to be complementary if they have separate technical elements that, when combined, can support and promote each other to increase their potential value and usefulness.



Introduction

Complementary Patent Identification



- **For countries**

  Break down the knowledge barriers between cross-domain technologies

  Address the challenges of product complexity and technological innovation uncertainty

- **In the course of industry development**

  Key industries such as information technology, communications and biomedicine

  Translate knowledge into technology and develop new products/services

- **In the case of R&D institutions**

  Difficult to master multiple fields and disciplines of technology at the same time

  Reduce the cost and risk of technology innovation direction and partner identification

mCube + Xsens

Introduction

- Current studies are usually based on the similarity between patents, ignoring patent complementarity.

➢ For example, there is a certain complementary relationship between the battery, electric control and motor of a car in terms of production process.

How to find complementary relationships like this and form a dataset

This research is planned to **construct a dataset** of complementary relationships annotated via hierarchical IPC classification numbers.

Introduction

Textual Content

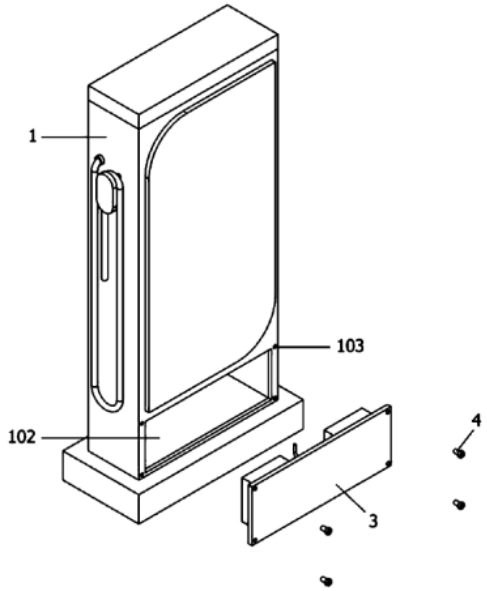
An essential part to acquire technical feature

- The present research relies primarily on patent classification or citation information, which may not accurately reflect the patent complementarity.

It is necessary **to incorporate their textual content** in order to delve deeper into their complementarity.

Title

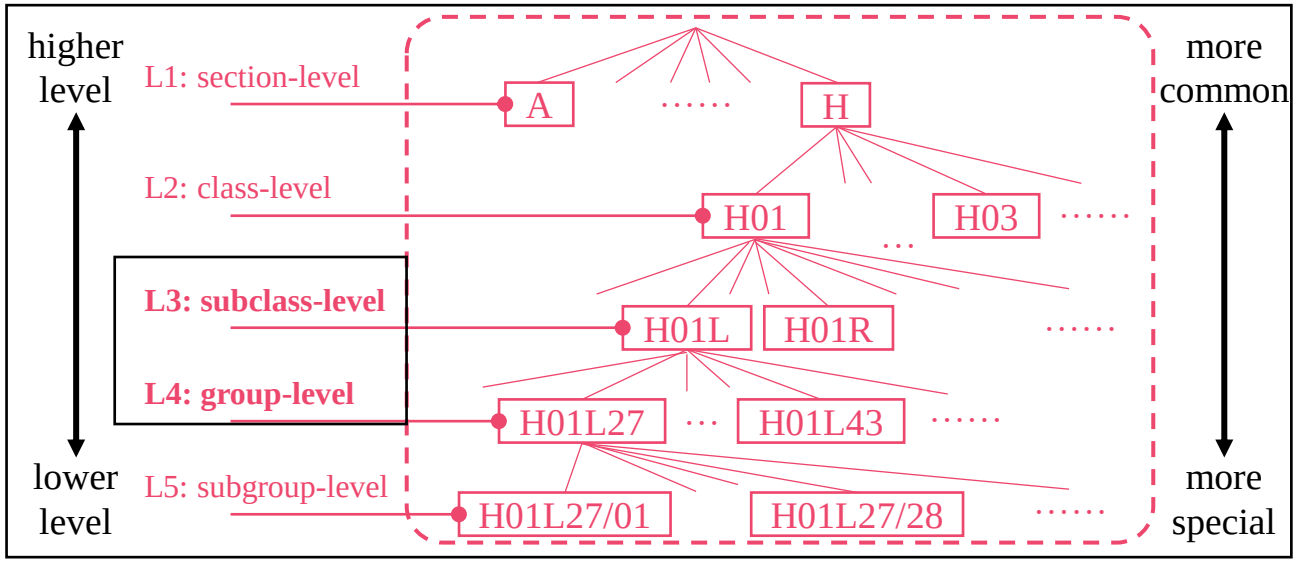
[EN] Charging pile with heat dissipation grid based on new energy automobile
[ZH] 一种基于新能源汽车的带散热格栅的充电桩



Abstract

[EN] The utility model provides a new energy automobile-based charging pile with a heat dissipation grid. The new energy automobile-based charging pile comprises a new energy automobile charging pile and apower supply connecting wire, a circular heat dissipation through hole communicated with the interior of the new energy automobile charging pile is formed in the middle of the adjacent bottom of thefront end face of the new energy automobile charging pile. Wind power generated after the cooling fan is started is blown to the inner wall of the new energy automobile charging pile; obstruction of the inner wall of the new energy automobile charging pile is avoided; wind power circulates in the new energy automobile charging pile; therefore, all electrical components in the new energy automobilecharging pile are contacted; therefore, heat generated by all the electrical components is driven to flow along with the electrical components. Under the condition that the two cooling fans continuously generate wind power, the wind power can burst through the notch structure formed in the center of the rubber sheet, so that the wind power carrying heat in the new energy automobile charging pileis discharged to the outside through the circular cooling through holes, and the effect of reducing the internal environment temperature of the new energy automobile charging pile is achieved.

Introduction



For example, studying the relationship between enterprises by the number of patents at the same group-lever or subgroup-level.

Certain research assesses complementarity according to the co-occurrence of patent classifications, which may not exactly indicate the degree of complementarity between patent.

This paper integrates deep learning techniques to establish a quantitative approach to identify complementary patents.

1. Introduction

2. Data and Method

3. Case Study

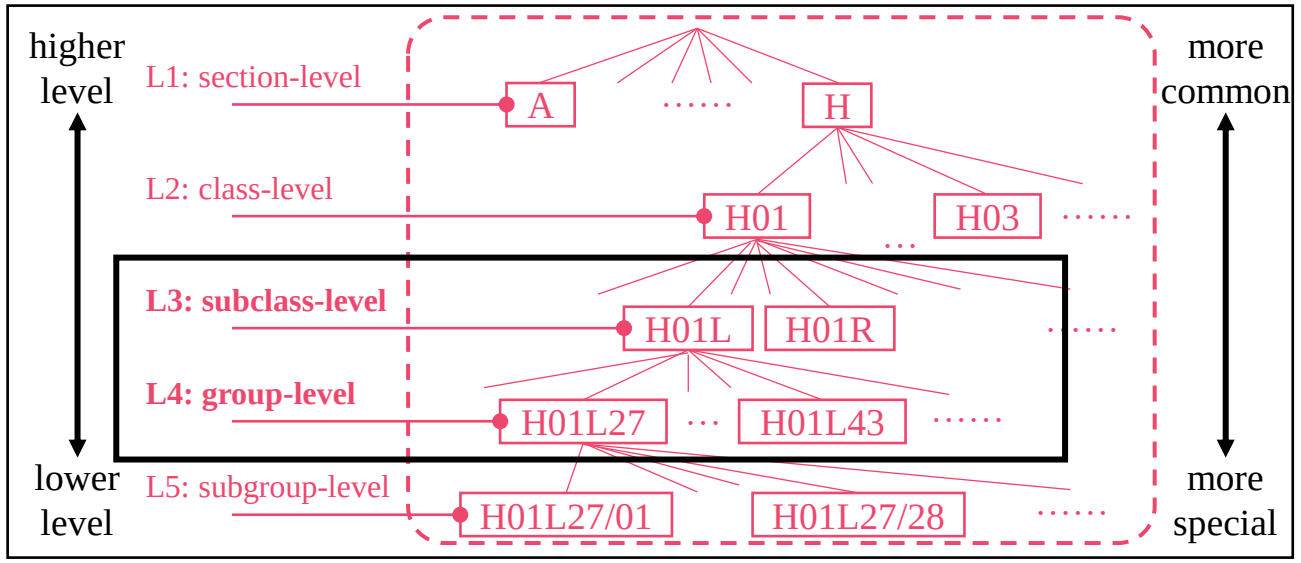
4. Conclusion and Future Works

Data and Method

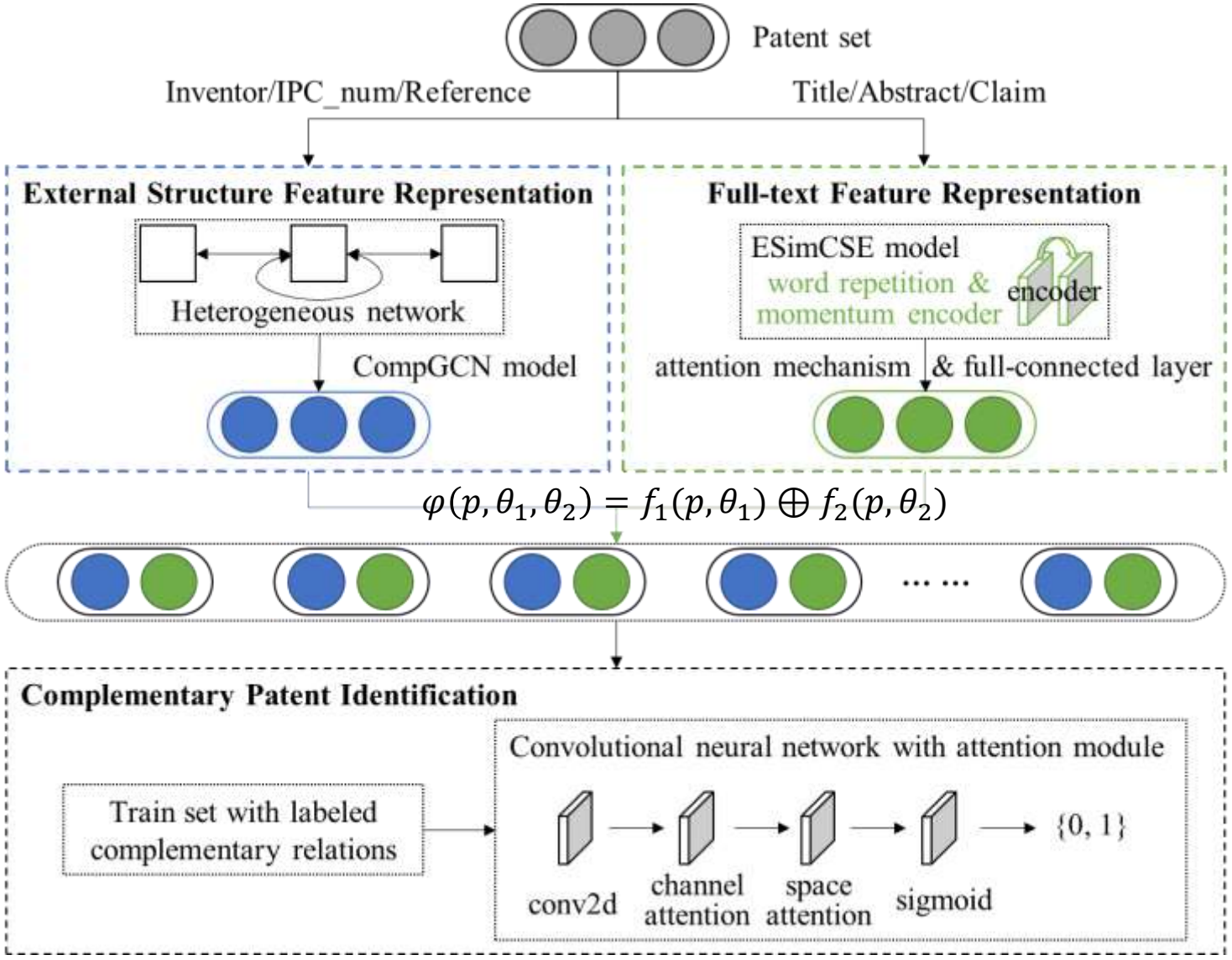
Data

Source: our study collected full-text patent data from USPTO for new energy vehicles published in 2022.

Complementary patent dataset: we determine whether a patent pair belongs to the same subclass-level but has distinct group-level according to the IPC classification numbers and assign a binary label of 0 or 1 to indicate their relationship.



Data and Method



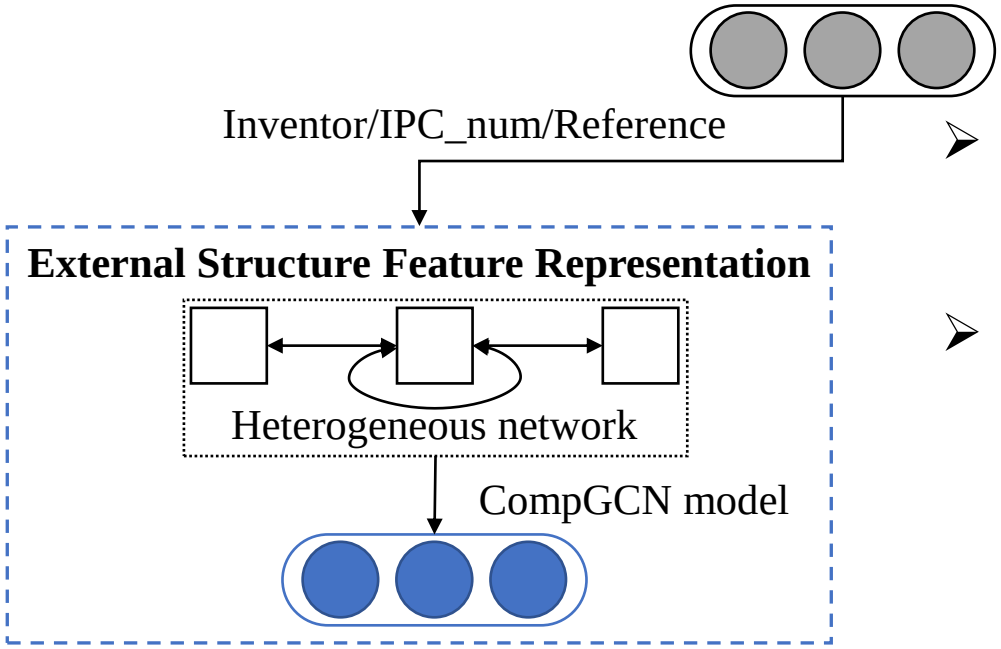
a patent set $P = \{p_1, p_2, \dots, p_m\}$

a triplet (p_i, r, p_j), $i, j \in (0, m)$

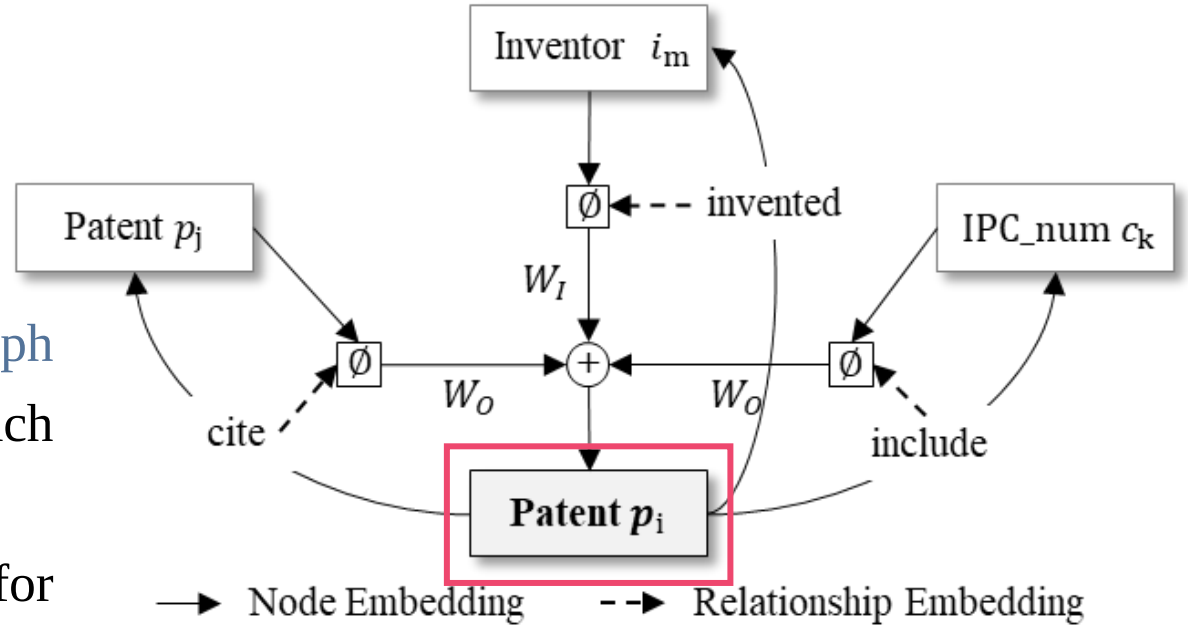
Goal: build a model to compute the complementary probability between patents p_i and p_j.

$$P(r = 1 | p_i, p_j) = F(\varphi(p_i, \theta_1, \theta_2), \varphi(p_j, \theta_1, \theta_2), \theta_3)$$

Data and Method



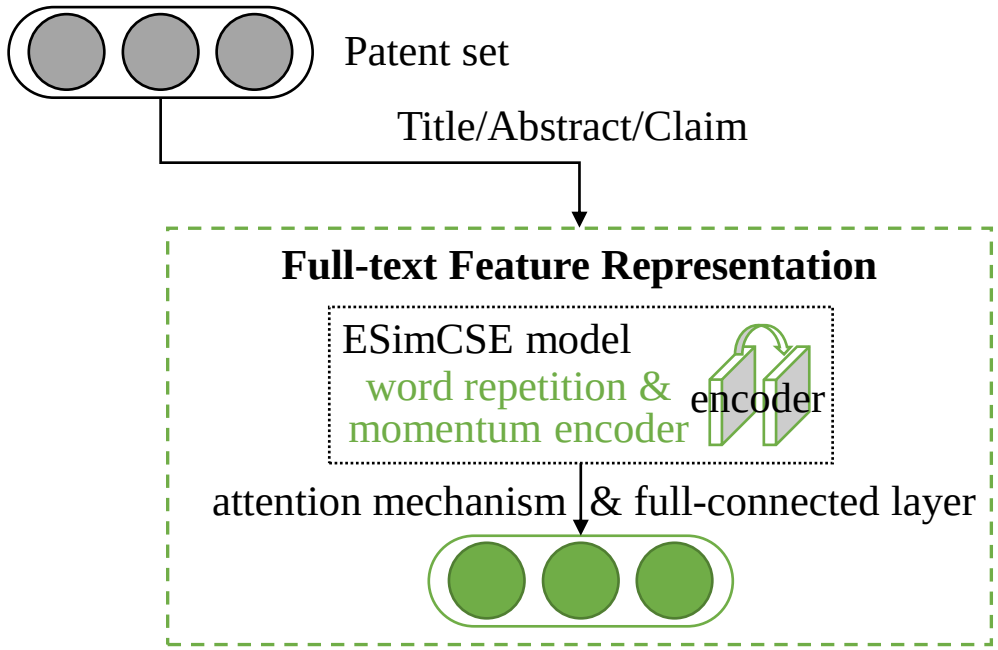
- 1. Collate the external structural items of patents, including inventors, IPC classification codes, and citation information.
- 2. Construct a patent heterogeneous network $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{O}, \mathcal{R})$



Node and edge aggregation process in the CompGCN model

- 3. Use the Composition-based Multi-Relational Graph Convolutional Networks model to extract rich information from the patent network
- 4. Obtain structural information representations for each patent.

Data and Method

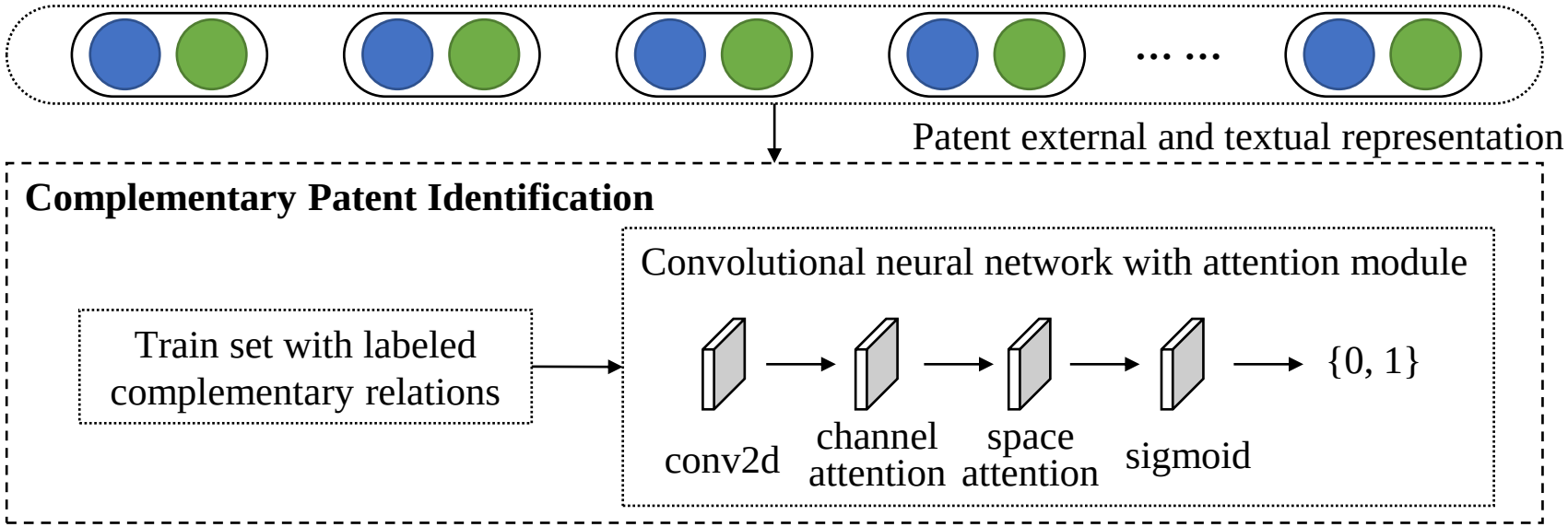


- 1. Use the Enhanced Simple Contrastive Learning of Sentence Embedding model to learn sentence vectors
- 2. Utilize attention mechanism to weigh the sentences and calculate the correlation between each dimension in the sentence vectors

$$o_l(s_a, s_b) = w^T \phi\left(W_{att}\left(l_a \odot l_b \odot f^l\right) + b\right) + c$$

(f^l denotes the unique heat vector representation of the lth pair of features, where the lth element is 1 and the rest elements are 0)

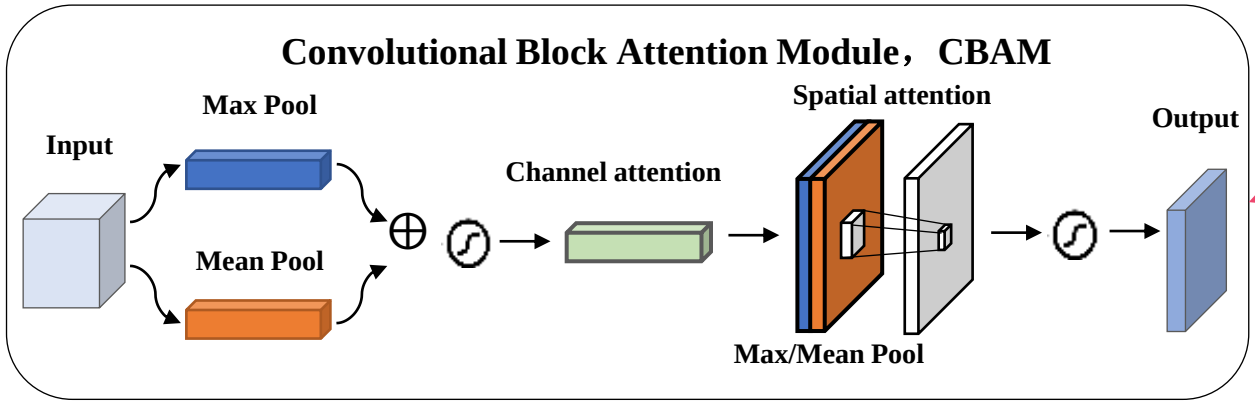
Data and Method



1. Treat complementary patent identification method as classification problem
2. Employ the CBAM module (Woo et al, 2018) to attend to channels and spaces that contain crucial information

3. Enhance the characteristics by a multi-layer neural network
4. Use a shifted sigmoid function on the fused features to calculate the final term

$$f_{p_i,p_j} = \sigma\left(\eta_d - c_{p_i,p_j}\right) = \frac{1}{1 + exp(c_{p_i,p_j} - \eta_d)}$$



Microgrid Expansion Planning Framework

Model Evaluation Metrics

- **Precision**： indicates the proportion of samples with positive predicted outcomes that are actually positive samples.
- **Recall**： indicates that the prediction result is the proportion of the actual number of positive samples in the positive sample to the number of positive samples in the full sample.
- **F1_Score**： a weighted average of precision and recall.

Compute the probabilities of complementary associations among patent pairs in the empirical dataset and expand the specific analysis.

1. Introduction

2. Data and Method

3. Case Study

4. Conclusion and Future Works

Case Study —— Contrast Analysis

Method	Precision	Recall	F1_Score
Structural	85.7	73.1	78.9
Textual	86.8	**76.9**	81.6
Structural& Textual	**90.3**	75.8	**82.4**

All the three methods yielded **precision above 85%**, demonstrating their ability to accurately identify complementary relationships with a high degree of matching and predict effectively.

The textual dimension method had **the highest recall rate**, indicating its proficiency in accurately identifying complementary relationships with a limited degree of matching.

Our proposed method attained **the highest F1 score**, indicating its superiority in identifying complementary relationships between patents.

Case Study —— Empirical Analysis

	Patent_nums	Probs
1	**US11431046 & US11522184**	**0.907**
2	**US11289723 & US11437666**	**0.901**
3	US11502350 & US11243260	0.889
4	US11322313 & US11502350	0.881
5	US11225166 & US11294551	0.876

- US11431046 proposes an energy storage device that can be used as an electrochemical battery containing positive and negative electrodes.
- US11522184 is a method of preparing positive active material.

The latter can be one of the directions of cooperation of the former patent, which can be used to **further improve the efficiency of energy storage**.

- US11289723 proposes a method to prevent icing of the exhaust system by controlling the fuel cell system.
- US11437666 proposes a battery module made of multiple side-by-side pouch-shaped battery modules.

For the former patent, optimizing the battery modules can further improve the performance of the fuel cell system and thus **better achieve its anti-icing goal**.

1. Introduction

2. Data and Method

3. Case Study

4. Conclusion and Future Works

Conclusion

1. Propose to analyze the relationship between patents from the perspective of complementarity and present a deep learning-based complementary patent identification method.

2. The proposed method takes only the basic patent information as input, with little pre-processing, and does not require laborious and expensive feature engineering.

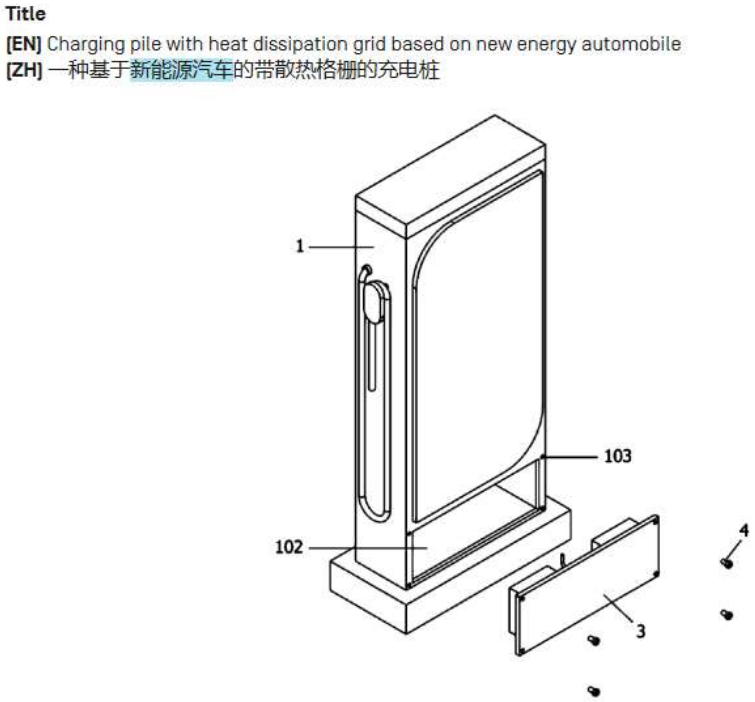
3. Using patents in the field of new energy vehicles to validate the method, and the importance of each part of the method is demonstrated by ablation experiments.

Future work

- Use **more complex network** and **text representation models** without compromising the simplicity of our approach.

- Use **patent image information** on top of this, such as hybrid models or mapping the external structure of patents, text and image information into the same embedding space, retrieving and matching each other.

Image



Reference

[1]. Zeng Deming & Zhou Tao. 2015. Research on the relationship between corporate knowledge infrastructure and technological innovation performance - a new perspective on the relationship dimension between knowledge elements. Science and Technology Management, 36: 80-88.

[2]. Douglas Henrique Milanez, Leandro Innocentini Lopes de Faria, Roniberto Morato do Amaral & José Angelo Rodrigues Gregolin. 2017. Claim-Based patent indicators: A novel approach to analyze patent content and monitor technological advances. World Patent Information, 50 (9): 64-72. doi: https://doi.org/10.1016/j.wpi.2017.08.008.

[3]. Marianna Makri, Michael A. Hitt & Peter J. Lane. 2010. Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. Strategic Management Journal, 31 (6): 602-628. doi: https://doi.org/10.1002/smj.829.

[4]. Shikhar Vashishth, Soumya Sanyal, Vikram Nitin & Partha Talukdar. 2020.
Composition-based multi-relational graph convolutional networks. Proceedings of the 8th International Conference on Learning Representations (ICLR 2020).
Addis Ababa, ETHIOPIA. doi: https://doi.org/10.48550/arXiv.1911.03082.

[5]. Xing Wu, Gao Chaochen, Zang Liangjun, Han Jizhong, Wang Zhongyuan & Songlin Hu. 2022. ESimCSE: Enhanced sample building method for contrastive learning of unsupervised sentence embedding. Proceedings of the 29th International Conference on Computational Linguistics. Gyeongju, KOREA: 3898–3907. doi: https://doi.org/10.48550/arXiv.2109.04380.

[6]. Sanghyun Woo, Jongchan Park, Joon-Young Lee & In So Kweon. 2018. CBAM: Convolutional Block Attention Module. Proceedings of the 15th European Conference (ECCV 2018). Munich, Germany. doi: https://doi.org/10.48550/arXiv.1807.06521.


Q & A
Thanks for listening