

Query Attribute Recommendation at Amazon Search

CHEN LUO, Amazon Search, USA

WILLIAM HEADDEN, Amazon Search, USA

NEELA AVUDAIAPPAN, Amazon Search, USA

HAOMING JIANG, Amazon Search, USA

TIANYU CAO, Amazon Search, USA

QINGYU YIN, Amazon Search, USA

YIFAN GAO, Amazon Search, USA

ZHENG LI, Amazon Search, USA

RAHUL GOUTAM, Amazon Search, USA

HAIYANG ZHANG, Amazon Search, USA

BING YIN, Amazon Search, USA

Query understanding models extract attributes from search queries, like color, product type, brand, etc. Search engines rely on these attributes for ranking, advertising, and recommendation, etc. However, product search queries are usually short, three or four words on average. This information shortage limits the search engine's power to provide high-quality services.

In this talk, we would like to share our year-long journey in solving the information shortage problem and introduce an end-to-end system for attribute recommendation at Amazon Search. We showcase how the system works and how the system contributes to the long-term user experience through offline and online experiments at Amazon Search. We hope this talk can inspire more follow-up works in understanding and improving attribute recommendations in product search.

CCS Concepts: • **Information systems** → **Information retrieval query processing**.

Additional Key Words and Phrases: Query Understanding, Product Search, Attribute Recommendation

ACM Reference Format:

Chen Luo, William Headden, Neela Avudaiappan, Haoming Jiang, Tianyu Cao, Qingyu Yin, Yifan Gao, Zheng Li, Rahul Goutam, Haiyang Zhang, and Bing Yin. 2022. Query Attribute Recommendation at Amazon Search. In *Sixteenth ACM Conference on Recommender Systems (RecSys '22)*, September 18–23, 2022, Seattle, WA, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3523227.3547395>

1 INTRODUCTION

Query understanding is the fundamental component of modern product search engines [2, 5]. Query understanding models extract attributes and intents from the customer query, such as color, product type, brand, etc. Then the search engine relies on these attributes for product ranking, recommendation, ads sourcing, etc [7]. However, customer queries are usually short. On average, they contain three to four words in product search engines [5]. The shortage of query information limits the power of the product search engine [6].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2022 Copyright held by the owner/author(s).

Manuscript submitted to ACM

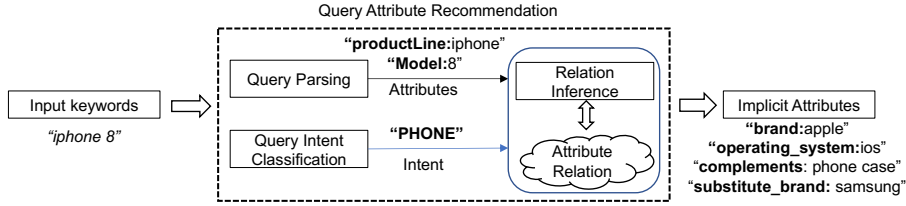


Fig. 1. The query attribute recommendation system. The query parsing model and query intent classification model first extract explicit attribute and intent from the query. Then the relation inference model extract the implicit attributes from an attribute relation graph.

Many works have been done in product search with limited query information. For example, [8] uses retrieved documents to generate the intent description for the query. Then the search engine uses the query intent description for improving exploratory search [8]. [1] improves the query expansion using external knowledge, such as wiki and wordnet. However, their improvement is limited due to the fundamental limitation of query side information that is relevant to product knowledge [3].

In this talk, we solve this problem differently from the above methods by directly enriching the query information through an attribute recommendation system at Amazon Search. We first introduce the three components of the attribute recommendation system: intent classification, explicit attribute parsing, implicit attribute recommendation. We then showcase how we use these additional attributes to improve ranking, ads, and query rewrite at Amazon.

2 QUERY ATTRIBUTE RECOMMENDATION

The attribute recommendation takes the customer input query and recommends the implicit attributes that do not explicitly appear in the search keywords. In Fig. 1, the input of the system is the customer query “iphone 8”, and the output of the system is a list of implicit attributes: ‘brand:apple’, ‘operating system:ios’, ‘complement:phone case’, and ‘substitute brand:samsung’. These implicit attributes enrich query information and provide additional features for product search applications such as product ranking, ads sourcing, query rewrite, etc.

The query attribute recommendation system contains three components: (1) a query intent classification model predicts the input query’s intent. (2) a query parsing model recognizes the explicit token’s attribute. Finally, (3) an implicit recommendation model that takes the output of query intent and query parsing to infer in an attribute relation graph and get the recommended implicit attribute.

3 QUERY INTENT CLASSIFICATION

In our system, we define the query intent as the product type intent of the query. For example, the intent of the query ‘iphone 8’ is ‘phone’, and the intent of the query ‘nike shoes’ is ‘shoe’. At Amazon, we have an ontology structure that defines all the product types on Amazon [3]. There are around 2000 to 3000 product types on the Amazon website. So, we frame the query intent prediction problem as a multi-label prediction problem in a multi-lingual, multi-country setting: given a query q and country id m , we predict whether each product type is relevant to the query.

We use past customer click behavior from the search logs to automatically label training data for our model. We aggregate each country, query, and the number of clicks on each product in 1 year’s data. Each product a is assigned a single label $L_a \in \mathcal{L}$ in the catalog. We construct labeled query examples as follows: For each unique country, query $x_i = (m_i, q_i)$, we define: $y_{it} = \frac{\sum_{L_a=t} N_{ia}}{\sum_a N_{ia}}$, where N_{ia} is the number of click for query x_i on product a .

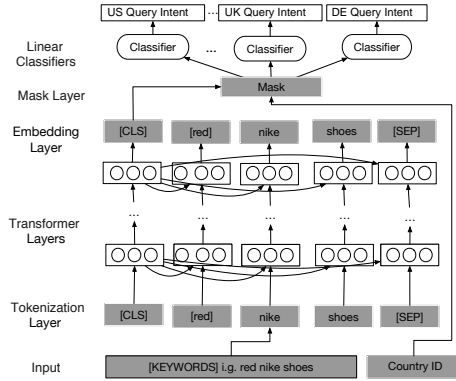


Fig. 2. The query intent classification model.

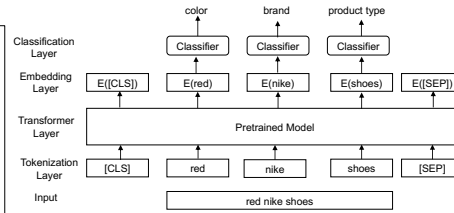


Fig. 3. The explicit query parsing model.

Our basic model builds on the transformer BERT model. After obtaining the embeddings for each token of the input query, we apply classification on top of the [CLS] output embedding. In this network structure, each country has a different label space, consisting of labels observed for products in that marketplace. We concatenate all the label spaces together, and each input country id m masks the corresponding labels for that country, as shown in Fig. 2.

4 EXPLICIT ATTRIBUTE PARSING

Query attribute parsing is the task of recognizing product attributes in search queries. For example, given a query ‘nike shoes’, the query attribute parsing identifies ‘nike’ as a brand token and ‘shoes’ as a product type. In our system, we design a transformer-based query parsing model, as shown in Fig. 3.

We use human-labeled data as our training data. The data contains 12 languages and more than 600K queries with ground truth tokens. The model structure of our query attribute parsing is in Fig. 3. We use a multilingual transformer model to handle the language agnostic input and use the transformer embeddings as the input of the final classification layer. Then each token obtains a classification label to denote what attribute this token belongs to.

5 IMPLICIT ATTRIBUTE RECOMMENDATION

To recommend implicit attributes given explicit attributes, we construct an attribute relation graph by learning from two data sources in Amazon: (1) product attribute data and (2) query attribute data.

The product attribute data is extracted from the catalog data. The catalog data contains all the products and the corresponding attributes. The sellers upload their products and input the corresponding attributes into our service. Then we automatically correct these attributes for each product to ensure the quality of the data [9]. For example, a product iPhone 8 has the attribute brand apple and product type phone.

The query attribute data is collected from customer search queries and their corresponding explicit attributes. We use the parsing model introduced in section 4 to obtain each query’s attributes.

There are over thousands of attributes in real product search services. In addition, there are hundreds of millions of queries every day. The size of the above data is prohibitive to process. Fortunately, the attributes from different product types or product intent don’t share much with each other. For example, the attribute *operating system* only appears on phones or laptop products. So, we split the data by product intent. The intent classification model obtains each query’s

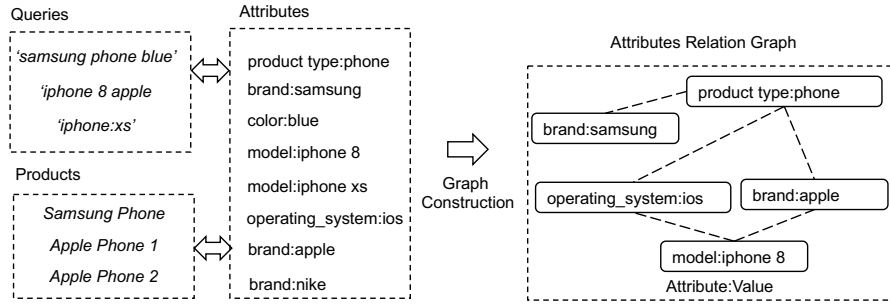


Fig. 4. The query implicit attribute recommendation

product intent, and each product’s intent is naturally obtained through catalog data. We also have a product knowledge graph [3] that assigns attributes to each product intent.

We use these two data to construct the attribute relation graph in Fig. 4. To get the relation between two attributes, we use GNN [4] model to generate embeddings for each attribute. Then we use these attributes embeddings to construct the relation between attributes.

During online inference, given the explicit query attributes and query intent, we use the attribute relation graph to recommend implicit attributes by doing graph random walk through the graph. We will introduce a detailed analysis of how we choose among different graph inference algorithms through the year-long investigation during the talk.

6 SYSTEM DEPLOYMENT AND IMPACT

We tested the system within Amazon for product ranking, product ads sourcing, and query rewrite, and proven to provide improved results.

- **Product Ranking** The product ranking model is a deep learning model that requires both features from the product and query. The implicit attribute provides a richer source of ranking features and improves the ranking performance that better matches the customer’s shopping intent.
- **Product Ads Sourcing** Amazon ads system relies on the search to source ads for customers. The implicit attribute improves the ads sourcing by increasing the choices of ads in line with customers’ shopping intent.
- **Query Rewrite** Query rewrite helps customers to reformulate the query that best matches customers’ shopping intent. The implicit attribute can provide richer choices of attributes to enhance the query rewrite process.

This talk will introduce more details on how we generate these implicit attributes and how we use them in Amazon to provide better services.

SPEAKER BIO

Chen Luo is an Applied Scientist at Amazon Search. He received his PhD from Rice University. His main research interests are in scalable machine learning with application in information retrieval and recommender systems. He publishes at ML and IR conferences and Journals such as WWW, KDD, SIGIR, and JMLR, and regularly serves as PCs for NeurIPS, ICML, KDD and WWW. He has given invited talks at many conferences, workshops and meetups.

REFERENCES

- [1] Hiteshwar Kumar Azad and Akshay Deepak. 2019. A new approach for query expansion using Wikipedia and WordNet. *Information sciences* 492 (2019), 147–163.
- [2] Yi Chang and Hongbo Deng. 2020. *Query understanding for search engines*. Springer.
- [3] Xin Luna Dong, Xiang He, Andrey Kan, Xian Li, Yan Liang, Jun Ma, Yifan Ethan Xu, Chenwei Zhang, Tong Zhao, Gabriel Blanco Saldana, et al. 2020. AutoKnow: Self-driving knowledge collection for products of thousands of types. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2724–2734.
- [4] Hanqing Lu, Yunwen Xu, Qingyu Yin, Tianyu Cao, Boris Aleksandrovsky, Yiwei Song, Xianlong Fan, and Bing Yin. 2021. Unsupervised Synonym Extraction for Document Enhancement in E-commerce Search. (2021).
- [5] Chen Luo, Vihan Lakshman, Anshumali Shrivastava, Tianyu Cao, Sreyashi Nag, Rahul Goutam, Hanqing Lu, Yiwei Song, and Bing Yin. 2022. ROSE: Robust Caches for Amazon Product Search. (2022).
- [6] Parikshit Sondhi, Mohit Sharma, Pranam Kolari, and ChengXiang Zhai. 2018. A Taxonomy of Queries for E-commerce Search. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 1245–1248.
- [7] Danqing Zhang, Zheng Li, Tianyu Cao, Chen Luo, Tony Wu, Hanqing Lu, Yiwei Song, Bing Yin, Tuo Zhao, and Qiang Yang. 2021. QUEACO: Borrowing Treasures from Weakly-labeled Behavior Data for Query Attribute Value Extraction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 4362–4372.
- [8] Ruqing Zhang, Jiafeng Guo, Yixing Fan, Yanyan Lan, and Xueqi Cheng. 2020. Query Understanding via Intent Description Generation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 1823–1832.
- [9] Wen Zhang, Yanbin Lu, Bella Dubrov, Zhi Xu, Shang Shang, and Emilio Maldonado. 2021. Deep Hierarchical Product Classification Based on Pre-Trained Multilingual Knowledge. (2021).