



Leveraging Search History for Improving Person-Job Fit

Yupeng Hou¹, Xingyu Pan², Wayne Xin Zhao^{1,✉}, Shuqing Bian²,
Yang Song³, Tao Zhang³, Ji-Rong Wen^{1,2}

1. Gaoling School of Artificial Intelligence, Renmin University of China

2. School of Information, Renmin University of China

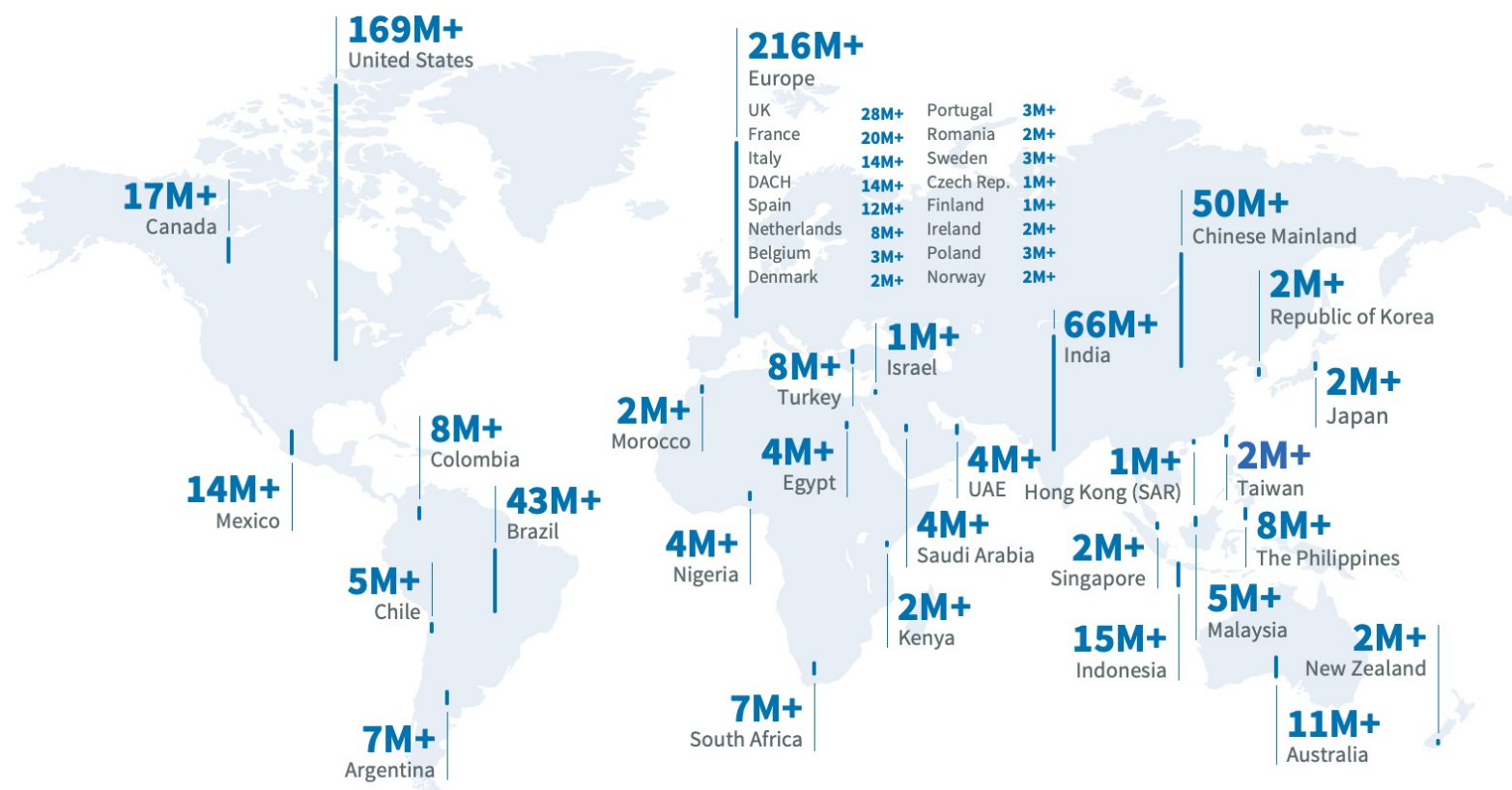
3. BOSS Zhipin



Background: online recruitment



600M Candidates
20M Jobs





Task: *Person-Job Fit*

Match qualified **candidates**
with suitable **job positions**

Job Requirement

1. Be familiar with **Python** language.
2. Have expertise in Machine Learning, Data Mining and Natural Language Processing.
3. Have experience with Large-scale data processing.
4. Have good communication skills and teamwork spirit.

Work Experience



Candidate A

I have participated in the students innovation team of university, as the team lead. I used the GBRT to predict the stock changes with the team members, and mainly focused on the model development.

Programming: **Python**



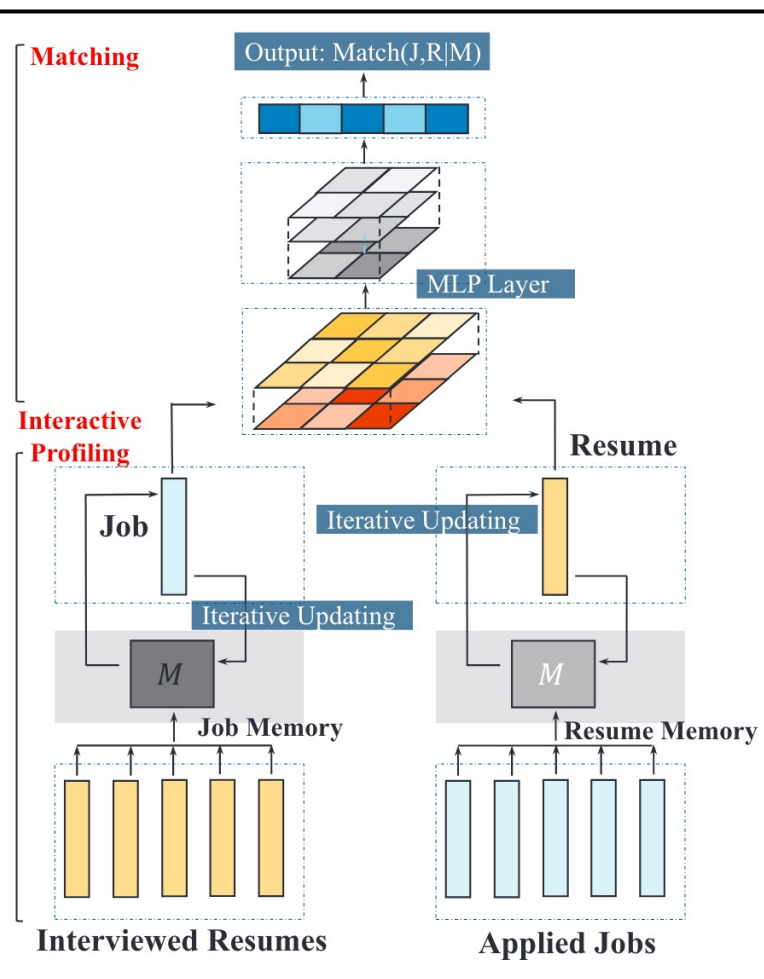
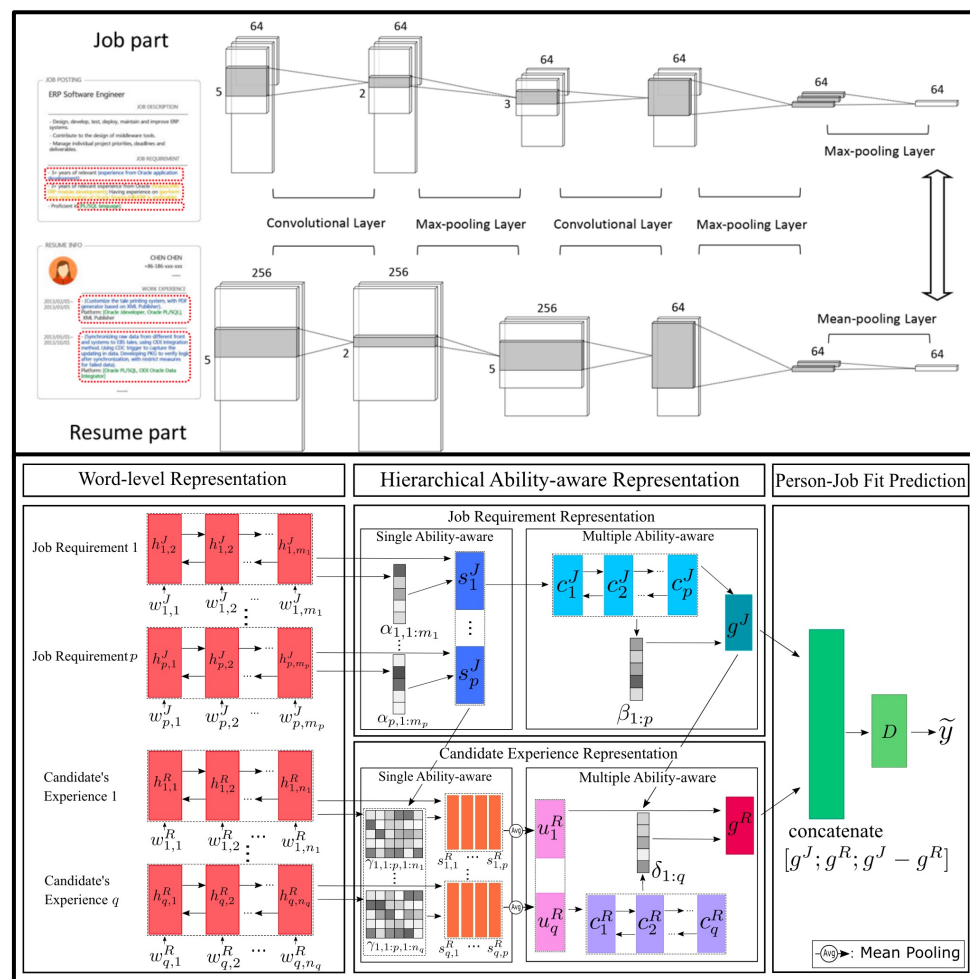
Candidate B

I have used Weibo textual data, combined with the characteristics of Weibo platform and the dissemination mechanism of false information. I have used the data mining technology and natural language processing technology to propose a Weibo credibility assessment algorithm.

Programming: **Python**

Red color: Programming Blue color: Machine Learning, Data Mining and Natural language Processing
Brown color: Large-scale data processing Green color: Communication and Team work

Related works



Text matching
(NN as text encoders)

Matching records from
recommendation scenario
as supervision signals

Zhu et al. "Person-Job Fit: Adapting the Right Talent for the Right Job with Joint Representation Learning." TMIS 2018.

Qin et al. "Enhancing person-job fit for talent recruitment: An ability-aware neural network approach." SIGIR 2018.



Yan et al. "Interview Choice Reveals Your Preference on the Market: To Improve Job-Resume Matching through Profiling Memories." KDD 2019.





Problem


- Two main scenarios, both exists interactions, text,

Recommended for you
Based on your profile and search history.

**Applied Scientist Intern (Intech Team, Beijing)**
Amazon
Beijing, Beijing, China
 2 company alumni work here
2 weeks ago



**Strategy Research Intern**
ByteDance
Beijing, Beijing, China (On-site)
 1 company alumnus works here
Promoted

Recommendation (~81%)

 data scientist China

Jobs | Date Posted | Experience Level | Company

Data scientist in China
40,856 results Set alert ☐

**Data Scientist & ML Operations Engineer - M365 IDEAs**
Microsoft
Changzhou-Wuxi-Suzhou Metropolitan Area
 113 alumni work here
Promoted • **8 applicants**

Search (~19%)



Problem

- Two main scenarios, both exists interactions, text,
- *Search*
 - has seldom been considered 😞
 - reflects important evidence for **job intention** of users 🤔



Problem

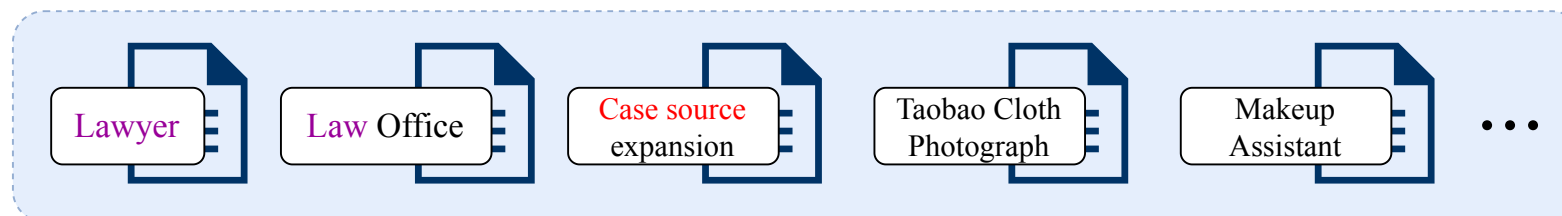
- Two main scenarios, both exists interactions, text,
- *Search*
 - has seldom been considered 😞
 - reflects important evidence for **job intention** of users 🤔
- Can we leverage these two sources simultaneously? 💡



Challenges

- Hard to capture job needs from search history

Search History

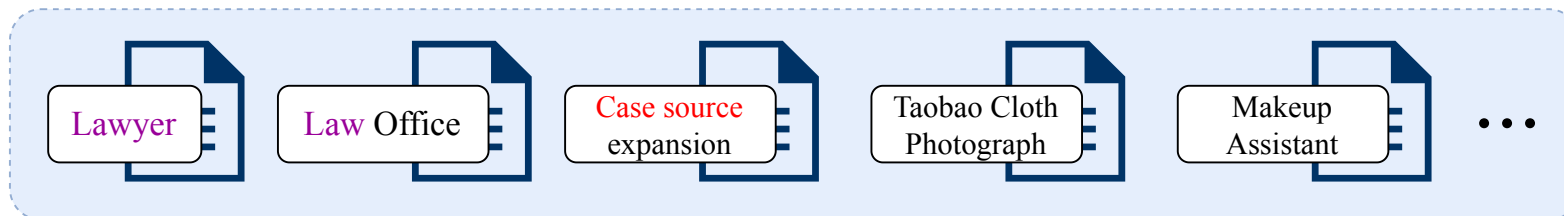




Challenges

- Hard to capture job needs from search history

Search History



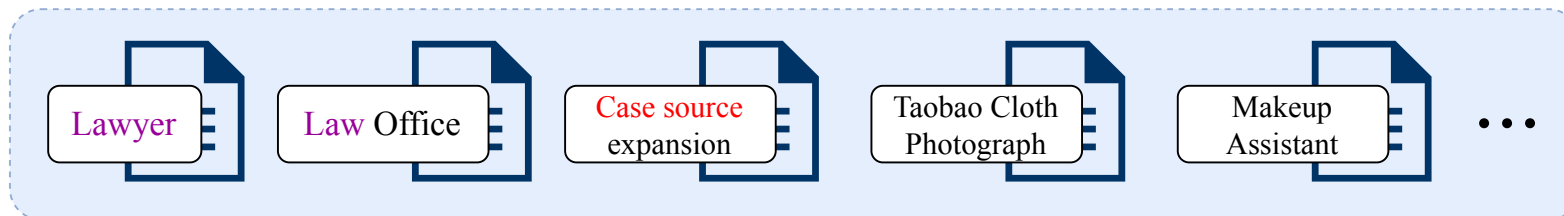
- **Data format:** list of short queries together with clicked jobs;



Challenges

- Hard to capture job needs from search history

Search History

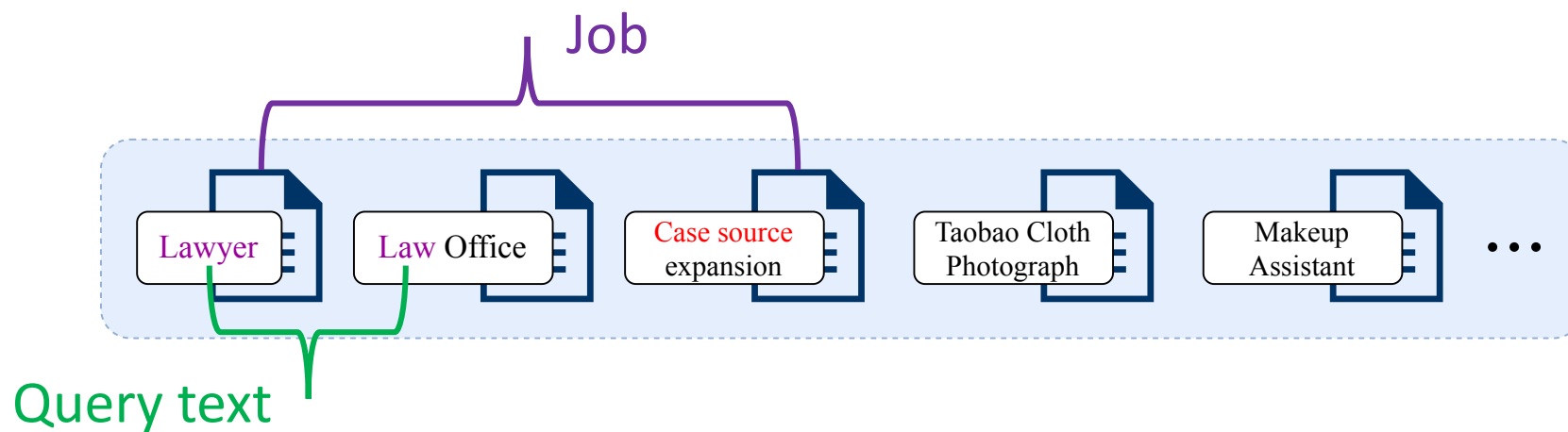


- **Data format**: list of short queries together with clicked jobs;
- **Redundant & Diverse**: how to find out underlying major intentions;



Idea

- 1. Cluster job intentions;
- 2. Assign attention weight on intentions based on:
 - Job similarity;
 - Query text similarity;





Problem Definition

- $u, j, \mathcal{H}_u \rightarrow y$
- Candidate: u
- Job: j
- Search history: $\mathcal{H}_u = \{\langle q_1, j_1 \rangle, \dots, \langle q_L, j_L \rangle\}$
<Query text, job ID>
- Matching or not: y



Solution- SHPJF

- Search History enhanced Person-Job Fit
- Text Matching Component
- Intention Modeling Component



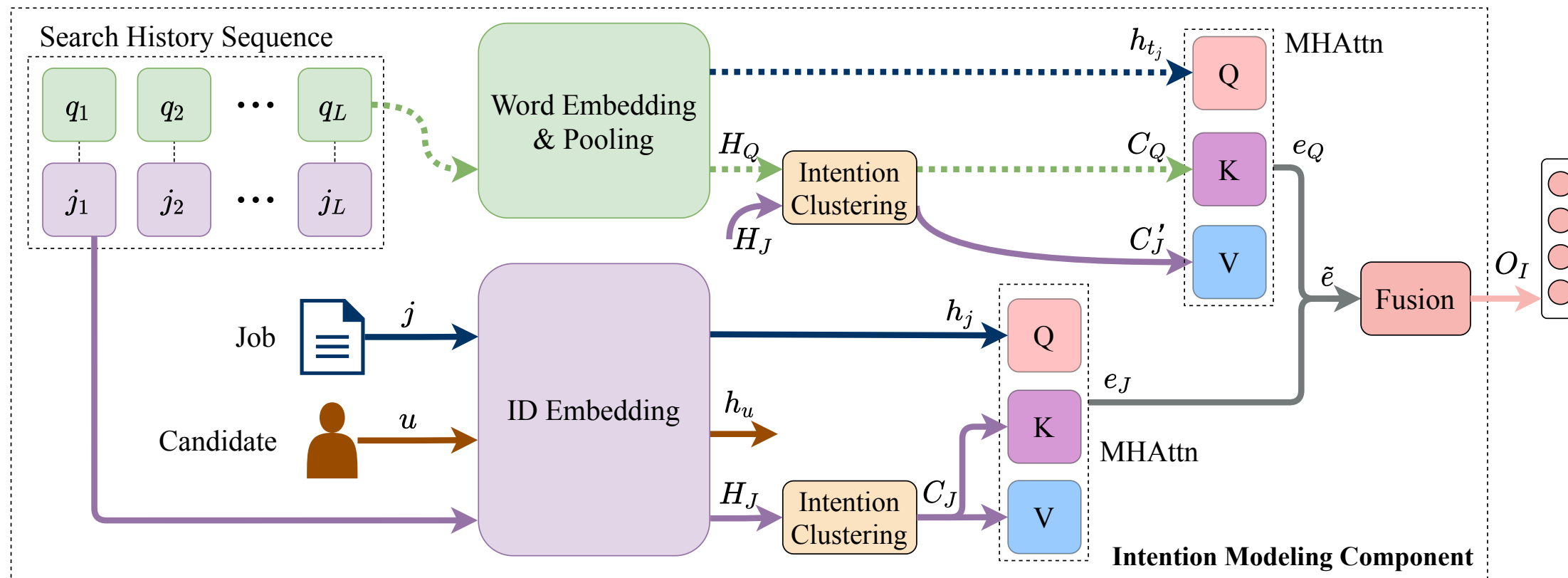
Solution – Text Matching



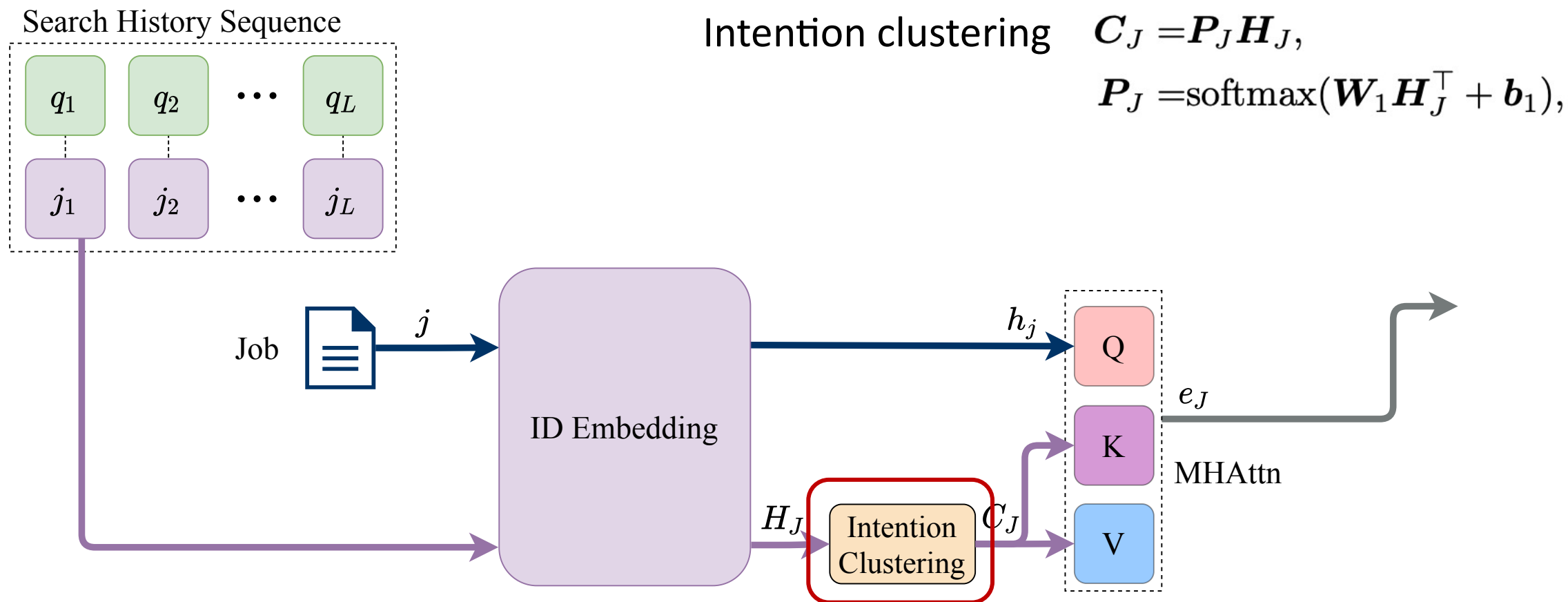
$$o_T = \text{BERT}([CLS]; r_u; [SEP]; t_j),$$

single-tower encoder

Solution – Intention Modeling

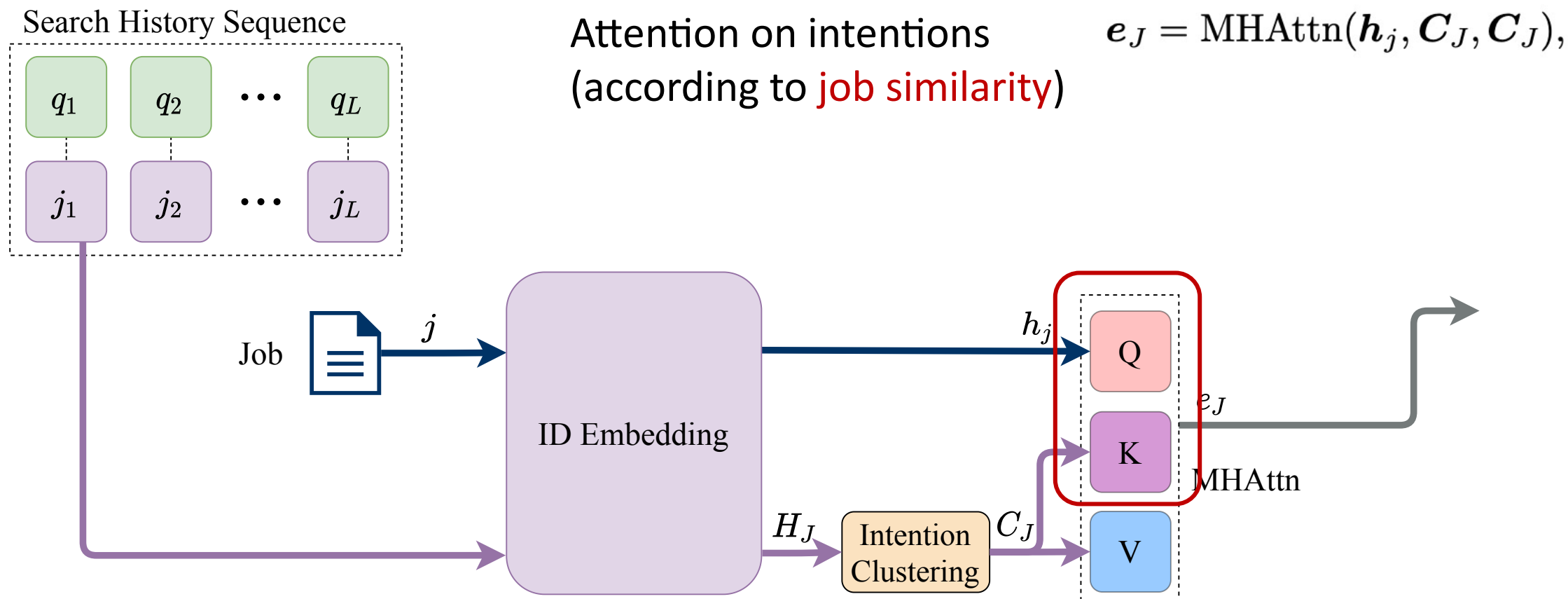


Solution – Intention Modeling (Job)

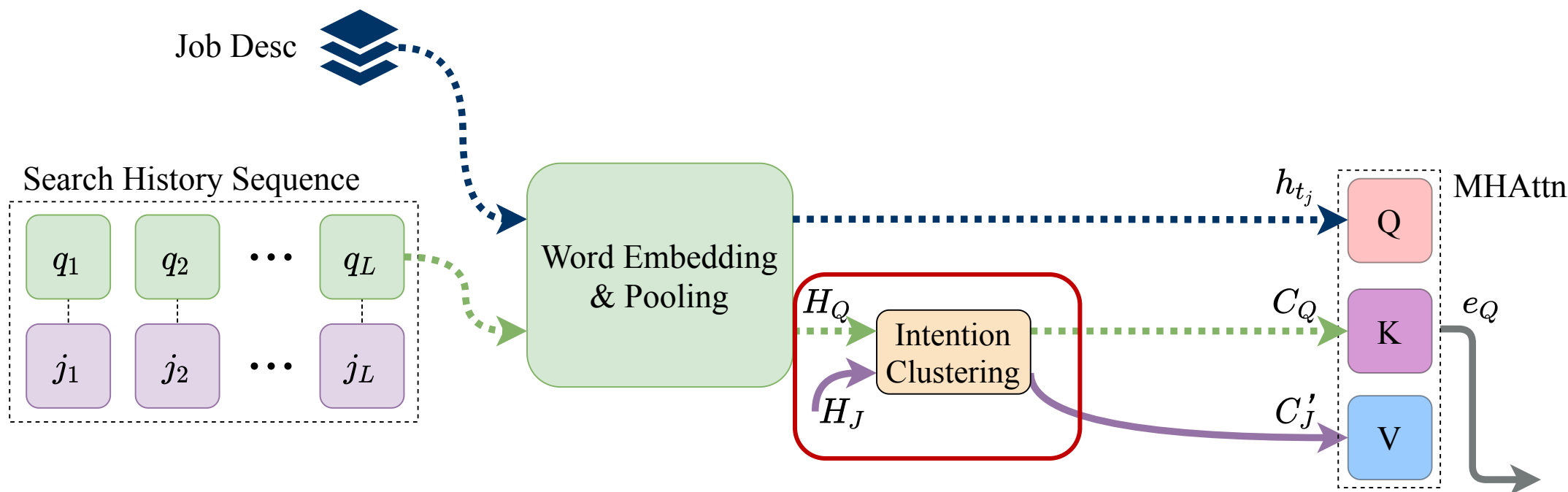




Solution – Intention Modeling (Job)



Solution – Intention Modeling (Text)



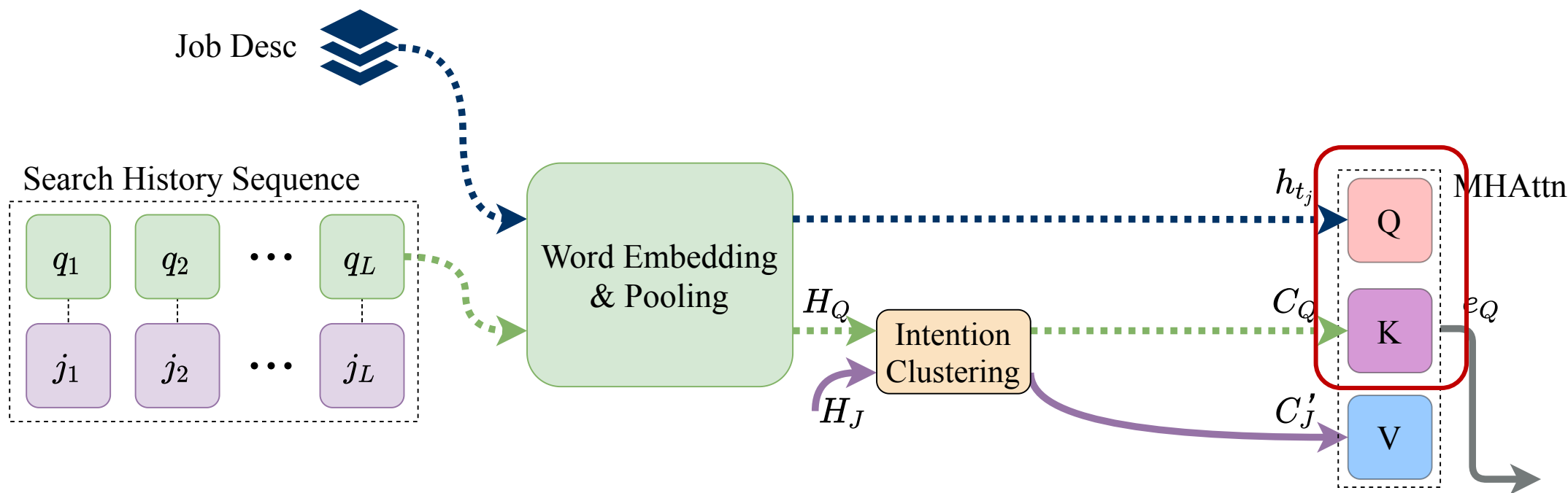
Intention clustering

$$C_Q = P_Q H_Q,$$

$$C'_J = P_Q H_J,$$

$$P_Q = \text{softmax}(\mathbf{W}_2 [\mathbf{H}_Q; \mathbf{H}_J]^\top + \mathbf{b}_2),$$

Solution – Intention Modeling (Text)



Attention on intentions
(according to **text similarity**)

$$e_Q = \text{MHAttn}(h_{t_j}, C_Q, C_J),$$

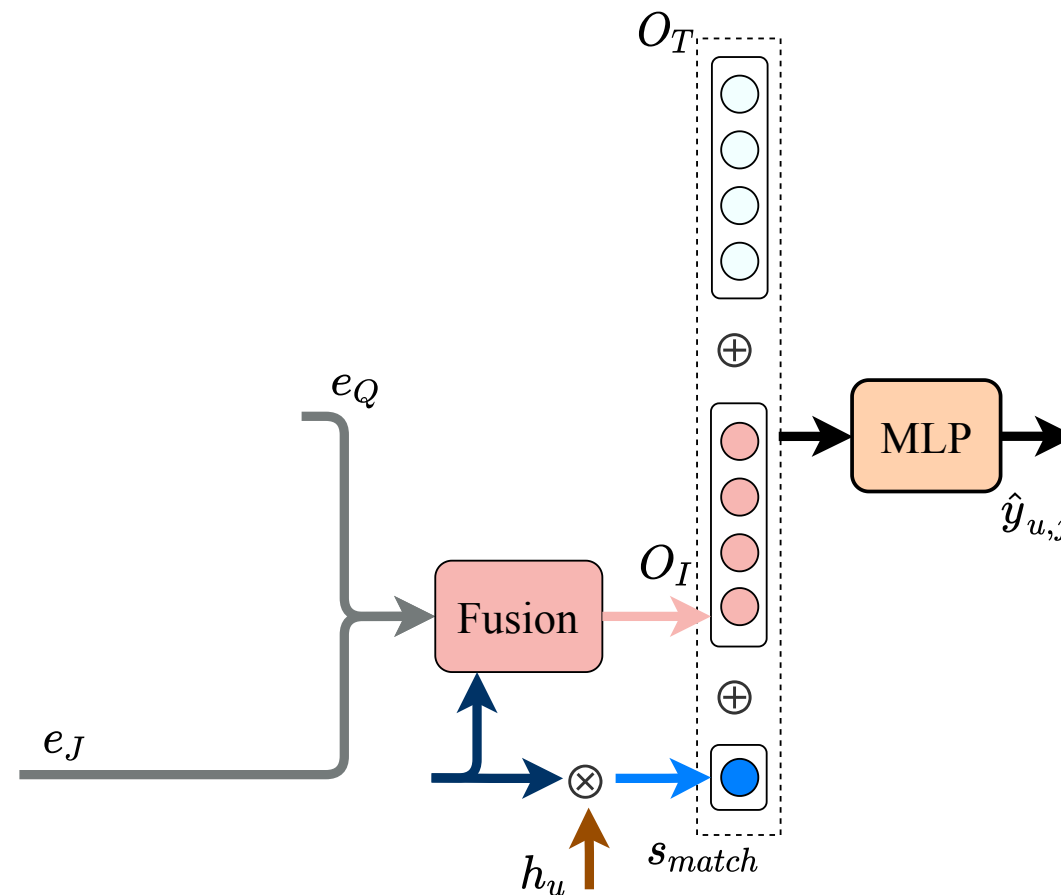


Solution – Fusion

$$\tilde{e} = \lambda e_J + (1 - \lambda) e_Q,$$

$$\mathbf{o}_I = \text{MLP}([\tilde{e}; \mathbf{h}_j; \tilde{e} - \mathbf{h}_j; \tilde{e} \circ \mathbf{h}_j]),$$

$$\hat{y}_{u,j} = \sigma(\text{MLP}([\mathbf{o}_T; \mathbf{o}_I; s_{match}])),$$





Experiments

- Dataset
 - Collect from BOSS Zhipin's real online log

#candidates	#jobs	#positive	#negative	\bar{L}	$\overline{ q }$
53,566	307,738	257,922	2,109,876	16.55	1.50



Experiments

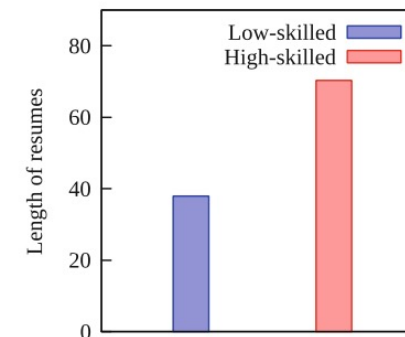
- Main results

Method	GAUC	R@1	R@5	MRR
PJFNN	0.5313	0.1412	0.5192	0.4025
BPJFNN	0.5343	0.1391	0.5217	0.4008
APJFNN	0.5323	0.1403	0.5185	0.4000
BERT	0.5449	0.1515	0.5297	0.4129
MV-CoN	0.5463	0.1554	0.5307	0.4165
SHPJF (ours)	0.5785	0.1630	0.5516	0.4297

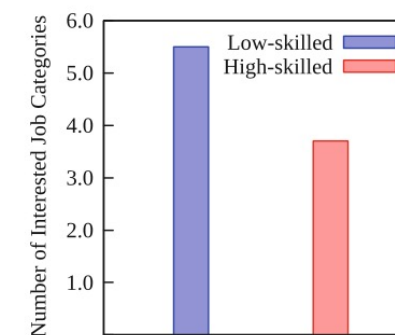


Experiments

- Evaluation in different skill groups
 - Short resumes
 - Multiple intentions



(a) Average resume length.



(b) Average #interested categories.

Groups	Low-skilled Candidates		High-skilled Candidates	
Method	GAUC	RelaImpr	GAUC	RelaImpr
PJFNN	0.5295	−4.53%	0.5318	−2.75%
BPJFNN	0.5399	+29.13%	0.5326	−0.31%
APJFNN	0.5309	0.00%	0.5327	0.00%
BERT	0.5381	+23.30%	0.5470	+43.73%
MV-CoN	0.5396	+28.16%	0.5484	+48.01%
SHPJF	0.5689	+122.98%	0.5814	+148.93%



Experiments


- Ablation Study

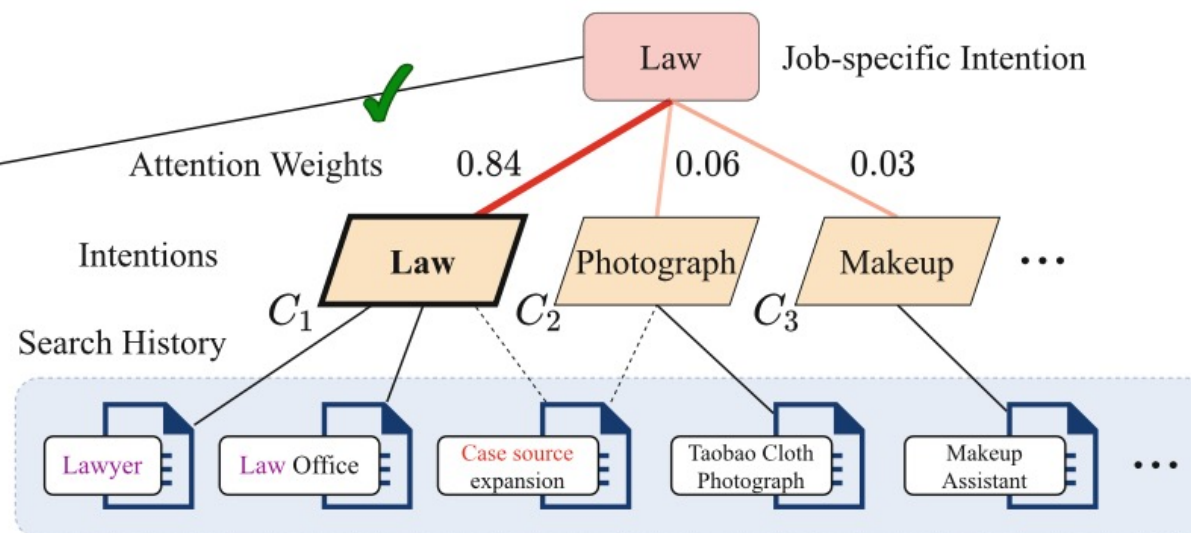
Variants	GAUC	R@1	R@5	MRR
BERT	0.5449	0.1515	0.5297	0.4129
BERT _{GRU}	0.5557	0.1546	0.5334	0.4196
BERT _{query}	0.5572	0.1558	0.5342	0.4201
SHPJF w/o Q	0.5697	0.1599	0.5456	0.4270
SHPJF w/o J	0.5715	0.1634	0.5456	0.4286
SHPJF w/o C	0.5738	0.1581	0.5443	0.4237
SHPJF	0.5785	0.1630	0.5516	0.4297

- Case Study

1. Responsible for the mediation work after the **copyright owner's** work is infringed.
2. The company provides the **source of the case**, the defendant, etc. Even freshmen can get started quickly.
3.

Proficient in Office softwares,
computer operation and usage of
printers.


Candidate
ID: ***6830





Conclusion

- SHPJF: Search History enhanced Person-Job Fit
 - Incorporating text/interactions from two channels: recommendation & search
 - Assign attention weight on intention representations
 - Text similarity
 - Job similarity
- Future Work
 - Side information from both channels
- <https://github.com/RUCAIBox/SHPJF>
- Welcome stars! 🌟