CheckList

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] see in 6
   (d) Did you describe the limitations of your work? [Yes] See Section 5.
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [Yes] See in Appendix A
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [Yes]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [Yes]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes]

A Detailed Dataset Description

A.1 Detailed Feature Description

In this subsection, we will provide a detailed description of the features of the large-scale web search session mentioned in Sec. 3.1. Types and explanations on page presentation features and user behaviors are shown in Fig. 8 and Fig. 9, respectively. One thing needed to be clarified is that, the displayed time is not necessary to be the summation of the displayed time on top, middle and bottom. The reason is that if the document displayed height is too large, for example, half of the screen, the display time will be recorded by both the top and middle of the screen.

Table 8: Rich Page Presentation Information in Baidu-ULTR

Presentation Info	Feature Type	Explanation
Query	Sequential desensitization tokens.	The user issued query.
Displayed Title	Sequential desensitization tokens.	The title of document.
Displayed Abstract	Sequential desensitization tokens.	A query-related brief introduction of document under the title.
MD5 of Document	String	A string for identifying the document.
Ranking Position	Discrete Number	The document's displaying order on the screen.
Multimedia Type	Discrete Number	The type of url, for example, advertisement, videos, maps.
SERP to Top	Discrete Number	The vertical pixels of the SERP to the top of the screen.
SERP Height	Discrete Number	The vertical pixels of SERP on the screen.

Table 9: Rich User Behaviors in Baidu-ULTR

User Behaviors	Feature Type	Explanation
Query Reformulation	Sequential desensitization tokens.	The subsequent queries issued by users under the same search goal. A session can have multiple queries.
Click	Discrete Number	whether user clicked the document.
First Click	Discrete Number	The identifier of users' first click in a query.
Skip	Discrete Number	whether user skipped the document on the screen.
Final Click	Discrete Number	The identifier of users' last click in a query session.
Slipoff Count	Discrete Number	The count of document being sliped off the screen.
Slipoff Count After Click	Discrete Number	The count of slipoff after user click the document.
Dwelling Time	Continuous Value	The length of time a user spends looking at a document after they've clicked a link on a SERP page, but before clicking back to the SERP results.
Displayed Time	Continuous Value	document display time on the screen.
Displayed Count	Discrete Number	The document's display count on the screen.
Displayed Time on Top Screen	Continuous Value	The document's display time on the top 1/3 of screen.
Displayed Count on Top Screen	Discrete Number	The document's display count on the top 1/3 of screen.
Displayed Time on Middle Screen	Continuous Value	The document's display time on the middle 1/3 of screen.
Displayed Count on Middle Screen	Discrete Number	The document's display count on the middle 1/3 of screen.
Displayed Time on Bottom Screen	Continuous Value	The document's display time on the bottom 1/3 of screen.
Displayed Count on Bottom Screen	Discrete Number	The document's display count on the bottom 1/3 of screen.
Last Click	Discrete Number	The identifier of users' last click in a query.
Reverse Display Count	Discrete Number	The document's display count of user view with a reverse browse order from bottom to the top.

B Further data analysis

In this section, we provide detailed data analysis results corresponding to the data analysis conclusion in Sec. 3.3.

We briefly mention the data analysis results again as follows:

1. Long-tail distribution appears in many user behaviors.
2. The logged search results on the high-frequency queries are more relevant than those on tail queries.
3. The most important user feedback, click, shows a strong correlation with both page presentation feature, position, and other user behaviors, displayed count.

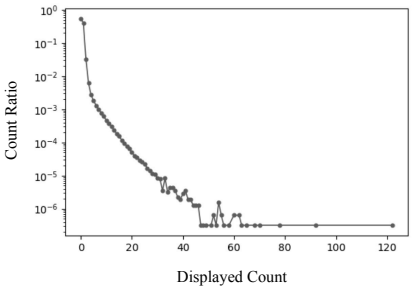

Figure 3: The distribution of displayed counts.

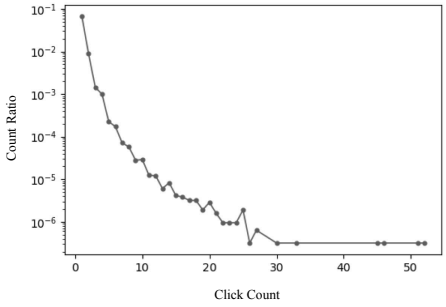

Figure 4: The distribution of click count.

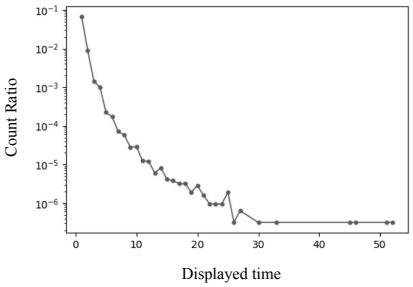

Figure 5: The distribution of displayed time.

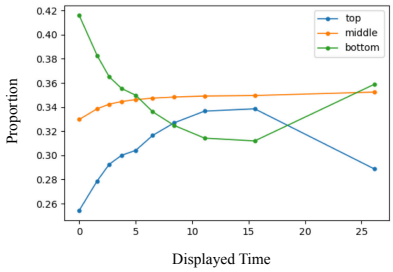

Figure 6: Displayed time ratio on different parts of the screen.

4. As the displayed time becomes longer, users are more likely to spend more time on the top of the screen.

Our supplementary focuses more on providing detailed results on conclusion 1 and conclusion 4. For conclusion 2 and conclusion 3, all analysis results are included in Sec. 3.3. For the conclusion 1 on long-tail distribution, we illustrate the distribution of displayed count, the displayed time, click count in Fig. 3, Fig. 5, Fig. 4, respectively. It is obvious that all those data distributions reveal a clear long tail phenomenon. Notice that for the continuous feature display time, we discretize the feature by splitting the interval 0.5 into the same bucket and calculate the count in each bucket.

For conclusion 4, the detailed data analysis results are shown in Fig. 6. The analysis procedure is as follows. We first remove documents with zero view time, which occupies a proportion of 54.7%. Then we ascendingly split data into 10 buckets according to the view time length, where each bucket has the same amount of documents. For each bucket, we calculate the proportion of document view time on the top, middle, bottom of the screen. For example, the proportion on the top of the screen is calculated by top displayed time / (top displayed time + middle displayed time + bottom displayed time). We can see that, as the displayed time becomes longer, users are more likely to spend more time on the top of the screen, from 25% to nearly 35%, while less time is spent on the bottom of the screen from more than 40% to 30%.