

Temporal Analysis of Demand and Supply of Topics on The Web

Masahiro Inoue
inoue@dl.soc.i.kyoto-u.ac.jp

Keishi Tajima
tajima@i.kyoto-u.ac.jp

Kyoto University

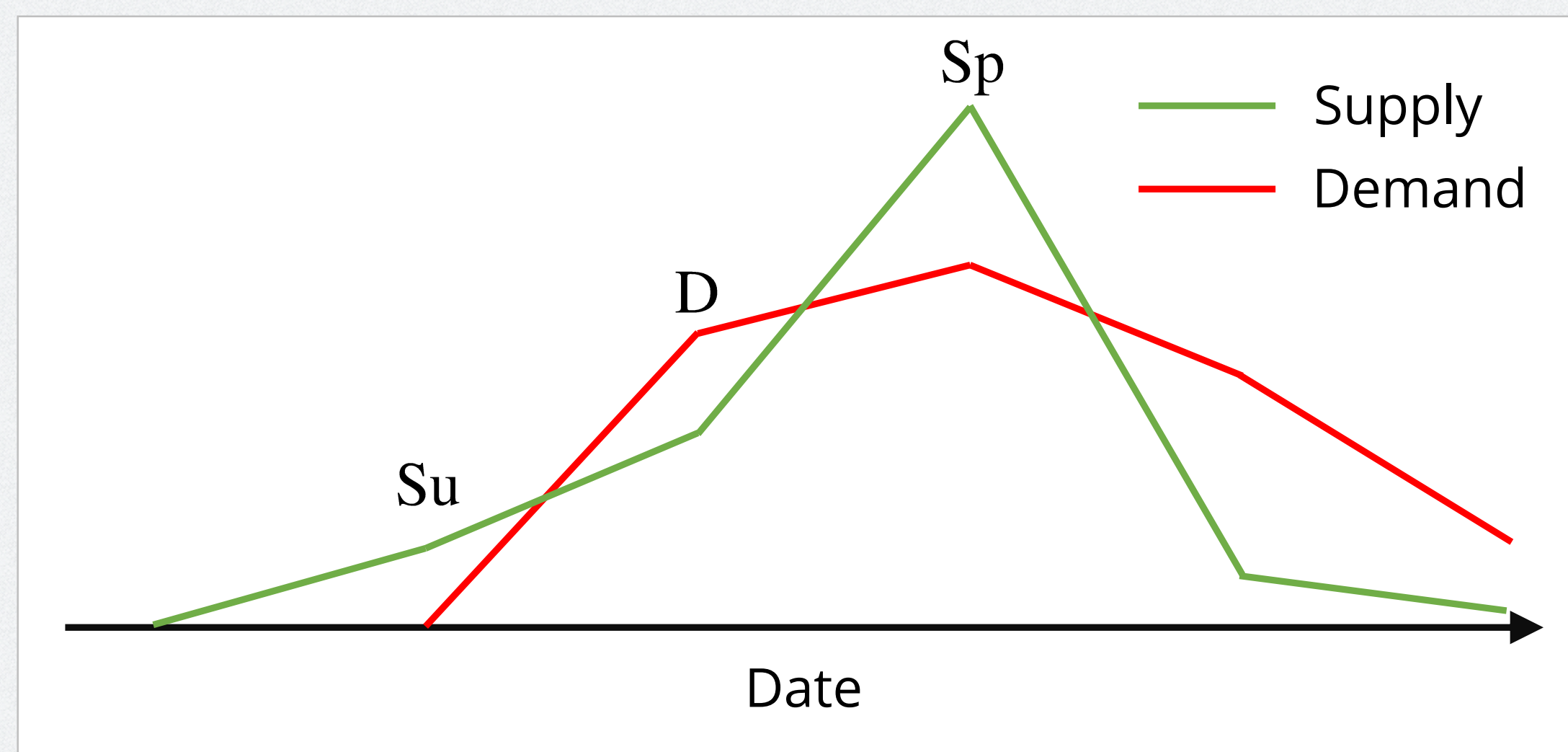
1. We analyze the supply and demand of topics on the web over time.

We analyze web topics and their respective queries by looking at the temporal relationship between the topics' significant dates of supply and demand.

- Su The supply start of the topic.
- D The demand start of the topic.
- Sp The supply peak of the topic.

We approximate:

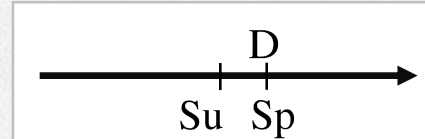
- D by the date when a query for the topic first appeared in Google Trends, and
- Sp and Su by the start and peak of the appearance dates of web pages in the query's search results.



2. We categorize web topics into six groups by the relative position of D.

A. Incidents/News

$$Su \leq D = Sp$$

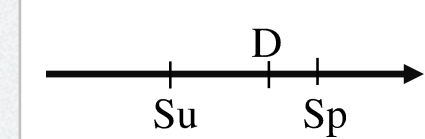


Typical topics in this group are news or incidents that occurred on the date $D (= Sp)$ and were unpredictable.

Example query:
Inside job documentary
(title of a movie awarded by the Oscar)

B. Prompt reports

$$Su \leq D = Sp - 1$$



This type of topic is first promptly reported with a basic summary of the results, and then published as a detailed article later within a day.

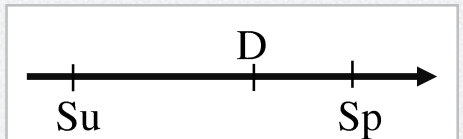
Example query:
Critics choice awards 2011

C. Big news

$$Su \leq D = Sp - (2 \text{ or } 3)$$

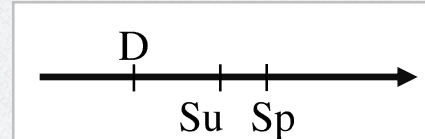
Topics in this group are typically events or discussions that last over the course of a few days.

Example query:
NFL draft
Arizona shooting



D. Real media topics

$$D < Su \leq Sp$$

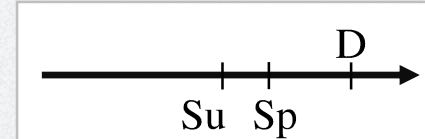


This type of topic is published first on other forms of media such as TV, and thus are not available on the web at the moment of the start of demand.

Example query:
Billy Walters (a gambler reported on TV)
Prader Willi syndrome (reported on TV)

E. Late noticed/Foreseeable

$$Su \leq Sp < D$$

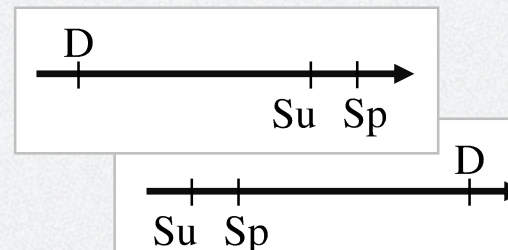


Two types of topics exist in this group. One is events that people noticed later than the time of occurrence, and the other is foreseeable events.

Example query:
Qwiki (an iPhone/iPad application)
Super bowl food (a planned event)

F. Famous or general nouns

$$|D - Sp| > 3$$



These kind of topics are mostly just general nouns or names of famous celebrities.

Example query:
Yosemite (location)
Kim Kardashian (an actress)

3. Conclusion

- We learned that we can categorize web topics by analyzing the temporal relationships of the significant points of supply and demand, even without understanding the topic directly.

- Furthermore, there is the interesting implication that within the groups A to C, topics in C has relatively earlier supply start dates (Su) than that of B, and those in B have relatively earlier start data (Su) than that of A. This indicates that upon obtaining the demand start date (D) and the supply start date (Su), it is possible to infer how far the supply peak date (Sp) appears in the future.