Stochastic Process for Analyzing Speech on the Web with Consideration of Media Mediation in Large-scale Broadcast Events in Japan

Yasuko Kawahata[1,2]
Rikkyo University, 3-34-1 Nishi-Ikebukuro,Toshima-ku,Tokyo[1]

Center for Spatial Information Science,
at the University of Tokyo, 5-1-5
Kashiwanoha, Kashiwa shi, Chiba[2]
Japan
kawahata@rikkyo.ac.jp

Abstract—In this paper, we will use the results of fitting with the sociophysical method regarding online media interaction in sports competitio etc (mainly this paper topics of rugby in 2016 to 2018 on broadcast) that has begun as a pioneer of large-scale broadcast events since 2000 in Japan.

Keywords—Online Social Networks · Sports Events · Blog · Stochastic processes

I. Introduction

This paper presents a case study on an approach to the distribution of opinions on web media in a large scale relay event using stochastic processes.(This study focuses on the sporting event Rugby, which was held in Japan.) The sports competitions mentioned in 2019 were selected. In this study, we obtained data from social media networks:SNS as a resource for data on the web in Japan. Fig.1 shows the total number of views of the wikipedia page in the major countries of SNS, which refers to social media networks, from 2015/7 to 2020/1/10. As can be seen from Fig.1, Japan has the highest number of web-based dictionary information views for SNS. It can be inferred that Japan has a wide range of communication on the Web, even when viewed globally. This suggests that Japan is a country that is particularly interested in the use of SNS as a tool for web-based communication, which has become popular since the late 2000s.

In addition, although Facebook is ranked in the top ten overall, Twitter tops the list except in English-speaking countries, indicating that there is a large number of users as well as interest in SNS and Twitter itself. In this context, this paper uses a dataset of Japanese SNSs to analyze the impact of the actual events on the tendency to mention large relay events in Japan to gain insights into the interaction between news/TV and viewers from a sociophysical perspective. Fig. 2-3 shows the number of mentions (2017/1-2020/1) of "events" (Fig.2) and "broadcasts" (Fig.3) in Japan for the major social networking sites (Twitter, 2ch and Blog). As can be seen, 'event' includes a variety of mentions of seasonal, traditional, annual and general events. It can be said that there are many different motivations for mentioning events in each medium and so on. In addition, the number of mentions in "live broadcasts" tends to be extremely skewed toward the two channels, especially in the spring and summer when live broadcasts of high school baseball and soccer programs are more prevalent. Although they are also broadcasted on TV, references to high school baseball and other games are especially close to 5ch (also called 2ch).

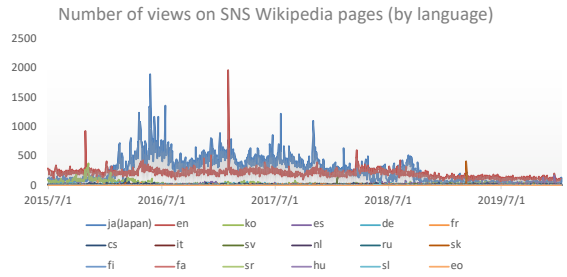

Fig 1 Number of views on SNS Wikipedia pages (By language ID)

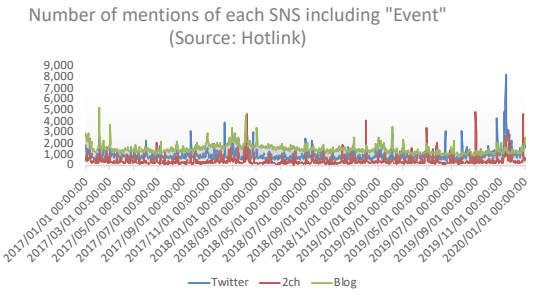

Fig 2. Number of mentions of each SNS including “Event” (Source: Hotlink)

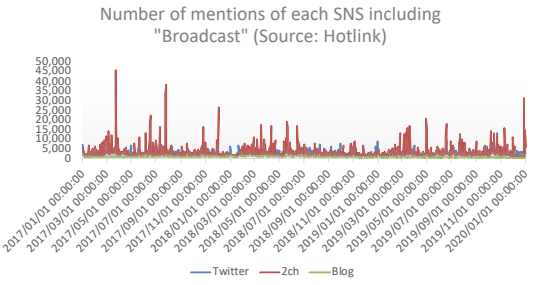

Fig 3. Number of mentions of each SNS including "Broadcast" (Source: Hotlink)

II. Dataset

In this research, we used 1 / 10-sampled Twitter data, 2ch data, a data set containing the keywords to be analyzed on major blogs in Japan, TV (reports per day) and News (reports per day). These datasets are provided by Hotlink. In addition, we used data sets provided by Elnet for the number

of references and topics in magazines and newspapers each month.

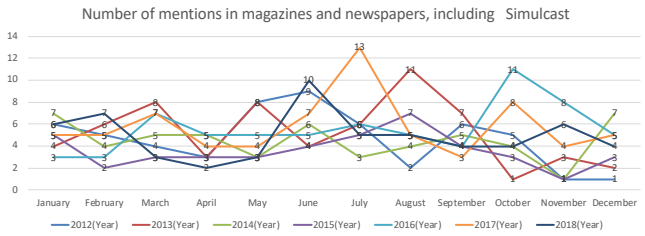

Fig 4 Number of mentions in magazines and newspapers, including Simulcast

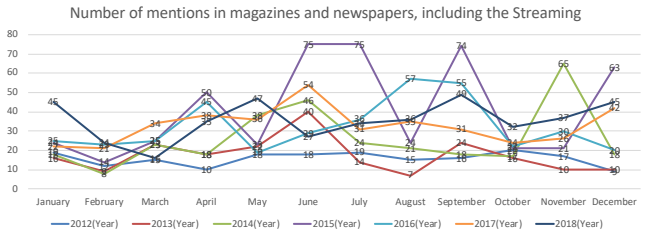

Fig 5 Number of mentions in magazines and newspapers, including the Streaming

As described above, it can be seen that references to "simultaneous relaying" are increasing year by year as a whole, but the bias in the number of references can be quantitatively estimated at the time when such technology is introduced and when it is estimated. We understand. It can be seen in the keyword co-occurrence word that discussions on relay technology in sports competitions such as the Olympic Games have been mentioned since 2018.There is also a similar tendency for references that are presumed to be focused on Internet broadcasting or music streaming, such as "Streaming", which will be issued as a keyword for large-scale broadcasting. In addition, it can be seen that both keywords tend to be mentioned in relation to special numbers at the end of the year. Fig. 6 shows the tendency to refer to "broadcast" in Japan since 2017. It can be seen that interest and live-states for the relay of the Diet in FY2017 were the highest. From 2018, it can be said that people's interest in sports (the Winter Olympics) and festival events such as coronation events have been changing.

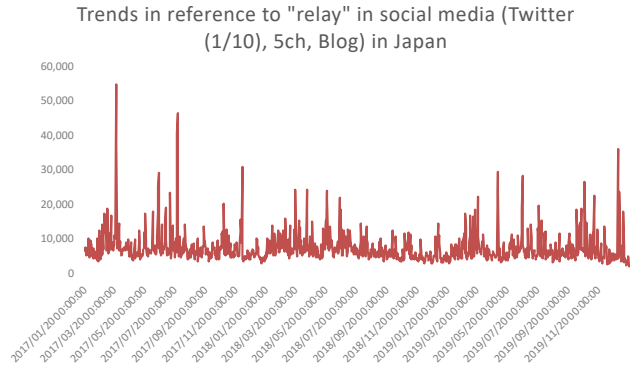

Fig 6 Trends in reference to "relay" in social media (Twitter (1/10), 5ch, Blog) in Japan

III. RESEARCH METHODS(FOCUS)

In this paper, we make a hypothesis that the transition of speech mentioned daily on social media is a phenomenon in one micro society, and use a mathematical model of social phenomena to consider those transitions from the viewpoint of social physics[1-2].

$$\frac{dI_i}{dt} = C_{adv}A(t) + \sum_j D_{ij}I_j(t) + \sum_j \sum_k P_{ijk}\, I_j(t)I_k(t) \qquad (1)$$

$$C_{adv}A(t) \Rightarrow \sum_i C_i A_i\,(t) \qquad (2)$$

(1) where Dij are the factor for the direct communication and Pijk is the factor for the indirect communication. The factor Cadv(News or TV) corresponds to the strength of the impression of the media or Concerts exposure for each advertising campaign. Because of the term of the indirect communication, this equation is a nonlinear equation. The factor Cadv corresponds to the strength of the impression of News or TV exposure for each advertising campaign. Because of the term of the indirect communication, this equation is a nonlinear equation. If we consider TV and News, we can solve the equation by including several effect of the advertisement in the following way (2).

IV. DISCUSSION

In this paper, we will use the results of fitting with the sociophysical method regarding online media interaction in rugby (a sports competition) that has begun as a pioneer of large-scale broadcast events since 2020. Rugby was written less often than baseball and soccer, as discussed in other sports-related references.

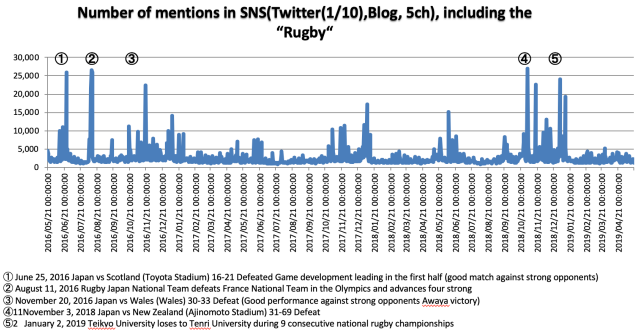

Fig 7 Number of mentions in SNS(Twitter(1/10),Blog, 5ch), including the "Rugby"

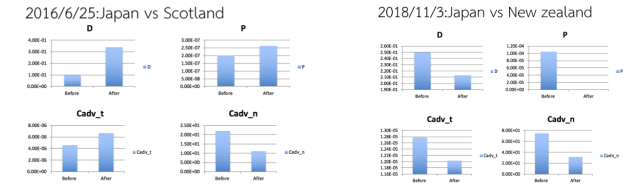

Fig 8 2016-2018 : Number of mentions of `` Rugby(in Japanese)" on the Twitter of Fitting
Parameta Result (D, P, Cadv_News, Cadv_TV) and Fitting Parameta Result

V. SUMMARY

All parameter values decreased before and after the 2018 game.It was assumed that the cause was a defeat in a large difference. It can be seen that the influence from the media is mutually great. In the 2016 game, although lost to the game, the D and P val- ues increased after the game. It can be seen that people are very interested in the game in Japan. Regardless of the outcome of the game, people's interests have been changing due to reports until 2018, but since the number of opportunities for relay and local support will increase throughout Japan, such as the World Cup, We think that it is necessary to consider later whether the direct (D) response that is not related to news or news will increase on the Web. From 2020-2025, it is expected that invitations will continue for cultural and sports events that also serve as international exchanges from large to small scales from the viewpoint of sustaina- ble operation of public facilities attached in the process.

VI. PURPOSE

The risks described above suggest that future case studies on how discourse spaces in online media are being shaped will be necessary. The Canon Institute for Global Studies (CIGS) meeting held in Tokyo, Japan (2020/11/12) also featured a number of discussions on public health using large-scale data. The following are some of the topics that were discussed.

In Japan, “COCOA”(Introduced in Japan from June 19, 2020) a smartphone app that uses Bluetooth functionality to track the physical contact opportunities of each individual, is now under scrutiny for providing information on security risks, such as coronaviruses, and managing risks in the event of large scale human flows. Similar applications are appearing in various countries, but their penetration rates vary considerably by country.

In Japan, the penetration rate was 14.4% as of October 2020.In the discussion, Japan, in particular, has the highest population density to ground area in the world (11th in the world (2019), 336(average) people/m², and the United States has 32.7(average) people/m²).From this perspective, tracking analysis using human flow data is actively being studied from a public health perspective. In particular. there are an expectation that idiosyncratic and other considerations can be obtained in densely populated cities such as Tokyo (47,006(average) people/m² in the case of New York City and 17,006(average) people/m² in the case of New York City) and Osaka).

We believe that it is very well suited for device-based studies related to infectious diseases and security incidents. However, in addition to “COCOA”, we are also actively working on demonstrations using proprietary applications at various university sites.

We think there will continue to be more discussions about how to better use information from online media, such as network tracking on devices, to protect individuals, and we think there will continue to be more research and opinions about actual security and infection risks using encrypted data and transmission models. . We believe that consideration can be obtained in the context of a large scale data analysis stream.

ACKNOWLEDGMENT

We thank for the support from the since 2017 to 2019 Scholars(Leading Initiative for Excellent Young Researchers (LEADER)), which provided support for this research in general. And also thanks for Toshimichi Wakabayashi of Department of Applied Mathematics and Physics for their research teams of a lots of supports (Tottori University) . And also thank the members of Rikkyo University's Department of Sociology and the Kawahata Seminar for their generous support in providing us with a new research environment from 2020.

REFERENCES

[1]A Ishii, H. Arakaki, N. Matsuda, S. Umemura, T. Urushidani, N. Yamagata, and N. Yoshida; The ‘hit’ phenomenon: a mathematical model of human dynamics interactions as a stochastic process, New Journal of Physics, 14, 063018 (2012).

[2]Kawahata, Yasuko, et al. "Analysis of mathematical model of hit phenomena stage actors of Japan." International Journal of Affective Engineering 13.1, pp.89-94(2014)