Effects of the inter-Korean summit on the presidents of N. Korea, S. Korea, and the USA evaluated using sediment analysis and Twitter corpus

KIM MIN SEOK [*], **KIM DONG HWAN**

[1, 2] St Paul's School, Sydney, Australia

Abstract

Aim: This study aims to determine if and how online evaluations of North Korean, South Korean, and American presidents affect their respective countries' public approval ratings.

Method: The HTML and Python programming languages were used to gather the tweets. SPSS and Microsoft Excel were used to analyze the data. An online lexicon and reasoning tool called Valence Aware Dictionary and Sediment Reasoner classified information as positive, negative, or neutral (VADER).

Findings: Independent Sample T-Test, Mann-Whitney U-Test, and Kruskal-Wallis Test were used to reject the null hypothesis of no significant difference between means. Despite popular belief, studies show that sediment score does not affect a show's performance offline.

Implications/Novel Contribution: Since using sediment analysis to prove approval ratings is new, this study makes a significant contribution. Researchers have used sediment analysis to assess the link between the online and offline worlds. To ensure a steady data stream for future studies, we can predict who will be president next with our Estimating the next president analysis. Trump's chances of being elected president in 2020 will be compared to Obama's performance in office.

Keywords: Sediment Analysis, Twitter Corpus, Inter-Korean Summit

INTRODUCTION

The world's political climate has deteriorated to the point where nations cooperate in a denuclearisation process. Denuclearization refers to a policy of not using nuclear weapons. N. Korea, however, is the only nuclear-armed state that was not a signatory. To put it another way, North Korea had been using nuclear weapons to threaten other nations and cause conflicts with others, including the United States and South Korea (Aristotle, 1438; Persson, 1995; Orthofer, 2016; Saahar, Sualman, Hashim, & Mohamed, 2017). Inter-Korean talks were held at Panmunjom on April 27th, 2018. Disarming Nuclear Weapons and Reuniting the World was the main focus. North Korea has also been using the media in its propaganda war. Trump's use of the word "lunatic" to describe North Korean leader Kim Jong-un quickly went viral. Speaking just before the Inter-Korean Summit, Donald Trump lauded the meeting and claimed he was looking forward to bettering relations between the two countries (Azhar, 2015; Kellner, 2018). The world is in disarray after North Korea made an abrupt shift in economic policy. The primary goal of this thesis is to investigate whether or not people's offline lives are impacted by access to online services (Frank, 2011; Louis, 2014; Parkin, Powell, & Matthews, 2017).

This paper analyzes user feedback on the three most prominent leaders at the Korean Summit: Donald Trump, Kim Jong-un, and Moon Jae-in. VADER, a computer program that assesses the feelings a sentence connotes, was used to compile the data on the keywords (Nair, 2017; Potty & Miller, 2013).

There are five major chunks of content in this paper. Following the introductory material is a comparison with other relevant studies. Technology and information resources utilized by the users are detailed in the third section. The null hypothesis is computed as well. The research findings, along with relevant statistics and visual aids, are presented in the following section. The information gathered on the subject is summarized in the final

---
[*]Corresponding author: Kim Min Seok
[†]Email: passenger020227@gmail.com

section.

Our study differs from others in that no one has ever used sediment analysis as evidence of approval ratings. Differentiating between online and offline contexts is the main focus of sediment analysis.

LITERATURE REVIEW

Twitter Corpus

Twitter, one of the world's most popular social networks on the internet has the potential for expanding business or website audience. Twitter has 326 million users (according to Omnicore, updated 10/26/18), and 500 million tweets per day (according to Omnicore, updated 6/24/18) which contains enough evaluation of issues every day. Twitter also reaches beyond Web browsers (Bardonova, Provaznik, Novakova, Sekora, & Svrcek, 2007). Twitter is able to be used to collect information in real-time. People use Twitter to post their messages online, practically sharing their thoughts. Collecting these types of information in real-time helps to analyze data more effectively than any other type of SNS.

Tweets were collected through direct searching from the web driver associated with HTML. Keywords were set in English, meaning only Tweets covered in English were collected. Approximately 2,000 tweets were collected each day. Trump overwhelmingly had more Tweets (about 1,400) compared to Kim and Moon (about 400, and 200 each). The biggest event was on April 27th, the day the inter-Korean summit took place (Sabine & Plumpton, 1985; Weimer & Weimer, 1993).

The biggest event that changed in sediment analysis was on the 27th, when the inter-Korean summit took place. Especially, Moon's positive evaluation percentage jumped nearly three times the past week. Unlike the others, Trump was not affected much since his positive evaluation went down about 2%. Kim had a steady position of positive evaluation after the 27th (Gondaliya & Gondaliya, 2012).

Hypothesis

The null and alternate hypothesis was related to the three major dates evaluated.

$$H0 : \mu1 = \mu2 = \mu3$$

HA: T□e means are not all equal.

μ was based on the day, data was collected. $\mu1$, $\mu2$, and $\mu3$ each stand for April 20th, April 27th, and March 4th.

The null hypothesis for Kim, Moon, and Trump is specified below.
The null hypothesis for Kim, Moon, and Trump is specified below.

$$H0 = \mu4 = \mu5 = \mu6$$

HA: T□e means are not all equal.

μ stands for each person, and the null hypothesis states that the results of the three people are equal. $\mu1$, $\mu2$, and $\mu3$ each stand for Moon Jae-in, Kim Jong-un, and Donald Trump.

METHODOLOGY

The research was based on figuring whether online evaluation can effect offline approval ratings. The research was based on the Inter Korean Summation.

Data was collected through Twitter by connection with Python coding. The data was gathered from both real-time and past data. The data was collected on coded formats and were turned to numeric data by coding. Then, a sensitivity analysis was conducted and visualized using SPSS and MS Excel.

The research was conducted with sediment analysis because it is thought to have been the most effective and crucial way to collect data according to human thoughts. Sediment analysis is words translated to scores, and the data base used for research was Twitter Corpus. Therefore, effective conversion of human thoughts to scores was conducted.

RESULTS AND DISCUSSION

Sediment Score

Table 1: Sedimental score of all dates

Type	St	Sediment Score (%)				
		4/13	4/20	4/27	5/4	5/11
Moon Jae-in	Pos	25.17	7.69	24.56	25.61	18.02
	Neg	17.14	9.62	12.28	8.54	14.41
	Neu	57.14	82.69	63.16	65.85	67.57
Kim Jong-un	Pos	30.53	15.00	21.77	27.51	26.89
	Neg	19.85	29.29	19.05	12.49	12.69
	Neu	49.62	55.71	59.18	60.00	60.42
Donald Trump	Pos	15.12	22.58	20.30	11.86	24.33
	Neg	22.71	12.27	16.45	33.19	15.23
	Neu	62.17	65.15	63.25	54.94	60.43

As a result, of the sediment score, the results show that Kim and Moon had an increase in positivity due to the effect of the inter-Korean summit, and Trump had a slight decrease. Especially, Moon's positive evaluation jumped approximately 3 times compared to the past week. Kim had a positivity score of 30.53% which was the highest of all ranks.

Statistical Analysis

Table 2: Levene's test and independent *t* test on Moon Jae-in

	Levene's Test		*t*-Test		
	F	Sig.	t	df	Sig.
Equal Variances not assumed	14.076	.000	4.926	511.251	.000

Levene's test was used to find the equality of variances. The independent sample *t*-test was used to find the equality of means. As a result of the tests, equal variances were no assumed ($F = 16$, $p = .079$).

Kruskal Wallis Test was used instead of one-way ANOVA in order to solve the problem related to the homogeneity of variances on dates.

Table 3: Levene's test and independent *t* test on Kim Jong-un

	Levene's Test		*t*-Test		
	F	Sig.	t	df	Sig.
Equal Variances not assumed	32.647	.000	1.524	5101.210	.128

Table 4: Levene's test and independent *t* test on Donald Trump

	Levene's Test		*t*-Test		
	F	Sig.	t	df	Sig.
Equal Variances not assumed	1.397	.237	5.086	1143.357	.000

Table 5: Kruskal-Wallis Test on three dates

Total N	Test Statistics	DF	Asymptotic Sig. (2-sided test)
5395	24.269	2	.000

According to the result of tests, null hypotheses were also rejected by the Kruskal-Wallis test in Table 4. There were statistically significant differences in three dates for Kim, Moon, and Trump. To examine the three people, the

Kruskal-Wallis test was performed. The three Kruskal-Wallis Tests prove practical significance between means. The means are all different according to the test which leads to practical significance in the data collected.

Table 6: Kruskal-Wallis test on Moon Jae-in

Total N	Test Statistics	DF	Asymptotic Sig. (2-sided test)
466	8.420	2	.015

Table 7: Kruskal-Wallis test on Kim Jong-un

Total N	Test Statistics	DF	Asymptotic Sig. (2-sided test)
1652	44.396	2	.000

Table 8: Kruskal-Wallis test on Donald Trump

Total N	Test Statistics	DF	Asymptotic Sig. (2-sided test)
3277	130.119	2	.000

According to the statistics results, it is shown that the means were not all equal providing evidence for practical significance. Consequently, differences of each artificial intelligence assistant were demonstrated in Table 6, 7, and 8.

Table 9: Kruskal-Wallis test on three People

Total N	Test Statistics	DF	Asymptotic Sig. (2-sided test)
111302	68.455	2	.000

There was a statistical difference among Kim, Moon, and Trump demonstrated by Kruskal-Wallis test in Table 9. Mann-Whitney test was applied in order to prove the difference between two independent samples.

Table 10: Mann-Whitney U Test Between Moon and Trump

Type	N	Mean Rank	Sum of Ranks
Moon	1057	4736.41	5006388.50
Kim	7840	4410.25	34576364.50
Total	10245		
	Score		Score
Mann-Whitney U	3839644.500	Z	-4.440
Wilicoxon W	34576364.500	Astmp. Sig(2-tailed)	.000

Mann-Whitney U Test was used to observe Moon and Kim. The rank of Moon was statistically higher than the rank of Kim. $U = 3839644$, $p = .000$, $r = -4$.

Table 11: Mann-Whitney U test between Kim and Trump

Type	N	Mean Rank	Sum of Ranks
Moon	2405	5476.32	13170549.00
Kim	7840	5014.62	39314620.80
Total	10245		
	Score		Score
Mann-Whitney U	8577866.000	Z	-7.595
Wilicoxon W	39314586.000	Astmp. Sig(2-tailed)	.000

Mann-Whitney U Test was used to observe the Kim and Trump. The rank of Moon was statistically bigger than the rank of Kim. $U = 8577866$, $p = .000$, $r = -7$.

Table 12: Mann-Whitney U test between Kim and Moon

Type	N	Mean Rank	Sum of Ranks
Moon	1057	1708.32	1805698.00
Kim	2405	1741.69	4188755.00
Total	3462		
	Score		Score
Mann-Whitney U	1246545.000	Z	-1.030
Wilicoxon W	1805698.000	Astmp. Sig(2-tailed)	.303

Mann-Whitney *U* Test was used to observe the Moon and Kim. The rank of Moon was statistically bigger than the rank of Kim. $U = 1246545$, $p = .303$, $r = -1$. Since the p value is not low, the Mann-Whitney *U* Test does not prove theoretical significance.

The Independent *t* Test proved that the means of data collected are all different. The Mann-Whitney *U* Test was conducted, but did not prove theoretical significance. The Kruskal-Wallis test was conducted and proved practical significance.

Discussion

The inter-Korean Summit was an important conference. North Korea provided a compromising point in nuclear weapons and an end to the Korean war. Also, it was the first event Trump actually praised an event linked to Kim Jong-un. The online evaluation of Trump had been significantly below 50-percent up until this event. However, Trump's evaluation struck positive for a while when linked to the inter-Korean Summit. Therefore, the author presumed that online evaluation could effect approval ratings (Miller, 1984; Sapiro & Rosenstone, 1999).

There were past researches that used sediment analysis to prove other options. Sediment analysis has never been applied to SNS corpus linked between both online and offline data.

The evaluation was conducted with Twitter Corpus and statistical evaluation supporting the visualization of data. There has been media wildfire about the election of Trump. The results of this research may prove how Trump had been using SNS and especially Twitter to interact with citizen of USA (Ismail, 2018).

This research may be expanded to show the connection between SNS and other materials. Also, Further data collection can lead to estimations on the next presidential election (ICPSR, 2004; Kelley, 2011; Shamir & Arian, 2018).

CONCLUSION, RECOMMENDATIONS AND IMPLICATIONS

This paper researched the capability of online evaluation and its affection towards the offline world. Tweets were collected using HTML and analyzed into three main stats (positive, negative, neutral). By VADER, a sediment dictionary, the raw data were transformed into scores. The changes in scores divided the accuracy of research. The Kruskal-Wallis test, Mann-Whitney *U* Test, Levene's Test, and Independent *t*-test was conducted to show the results of accurate values.

The difference between positive, and negative ratings was identified by statistical tests. The null hypothesis of Kim, Moon, and Trump was rejected and the alternative hypothesis was accepted. As a result, the author was able to conclude that online evaluation does not affect offline approval ratings.

There were some limitations to our research about the gathering of data. First, not all of the data online was consumed. There were limits to gathering the data. Also, only data in English format was collected. Other diverse types of data could have resulted in a more accurate conclusion. For future research, Estimation on the next president analysis will be capable of data is collected at stable rate. Comparison with the past president, Barack Obama will prove the possibility of Trump winning the next election which will be in November, 2020.

REFERENCES

Aristotle. (1438). *Politics*. New York, NY: Macmillan and Co. Limited.

Azhar, M. (2015). The concept of religious democracy as a new political philosophy for countries with moslem predominant. *Journal of Advances in Humanities and Social Sciences*, *1*(1), 19-28. doi:https://doi.org/10.20474/jahss1.1.3

Bardonova, J., Provaznik, I., Novakova, M., Sekora, J., & Svrcek, M. (2007). Statistical analysis in complex-valued wavelet analysis of voltage-sensitive dye mapping. In *2007 Computers in Cardiology,* New York, NY.

Frank, R. (2011). *North Korea in 2010: Domestic developments and the economy*. Hague, Netherlands: Brill.

Gondaliya, P. N., & Gondaliya, P. N. (2012). Sediment observation and its analysis. *Paripex-Indian Journal of Research*, *2*(2), 142-144. doi:https://doi.org/10.15373/22501991/feb2013/48

ICPSR. (2004). *American national election study, 2004: Panel study.* Retrieved from https://bit.ly/31cNOQ8 (Accessed on 1 August, 2017)

Ismail, I. (2018). *Election supervisory committees authority on election monitoring.* Retrieved from https://bit.ly/33JZiwj (Accessed on 15 August, 2019)

Kelley, J. (2011). *Data on international election monitoring: Three global datasets on election quality, election events and international election observation*. Abingdon, UK: Routledge.

Kellner, D. (2018). Donald Trump and the war on the media: From election 16 into the Trump presidency. In Robert, E. G. J (Ed.), *the trump presidency, journalism, and democracy*. Abingdon, UK: Routledge.

Louis, F. (2014). *Two election and removal of the president: The law of the executive branch*. New York, NY: Sage Publication.

Miller, W. E. (1984). *American national election studies, 1982: Post-election survey file.* Retrieved from https://bit.ly/35zkxSZ (Access on 15 August, 2017)

Nair, B. B. (2017). The politics of tourism representations: Discourse analysis of British travel, brochures about incredible India. *Journal of Advanced Research in Social Sciences and Humanities*, *2*(4), 230-236. doi:https://doi.org/10.26500/jarssh-02-2017-0401

Orthofer, M. (2016). *The complete review guide to contemporary world fiction*. Columbia, Portland, OR: Columbia University Press.

Parkin, M., Powell, M., & Matthews, K. (2017). *Economics*. Harlow, UK: Pearson.

Persson, T. (1995). *Politics*. Cambridge, MA: MIT Press.

Potty, G. R., & Miller, J. H. (2013). Time frequency analysis techniques for long range sediment tomography. In *Proceedings of Meetings on Acoustics,* California, CA.

Saahar, S., Sualman, I., Hashim, M. A., & Mohamed, W. A. W. (2017). Political advertising and young voters information needs. *International Journal of Humanities, Arts and Social Sciences*, *3*(5), 215-222. doi:https://doi.org/10.20469/ijhss.3.20004-5

Sabine, P., & Plumpton, C. (1985). *Statistics*. Basingstoke, UK: Macmillan Education.

Sapiro, V., & Rosenstone, S. J. (1999). *American national election study, 1998: Post-election survey* (Technical report). Inter-university Consortium for Political and Social Research, Michigan, MI.

Shamir, M., & Arian, A. (2018). *Context and comparison in election research: The Israel national election study*. Abingdon, UK: Routledge.

Weimer, R. C., & Weimer, R. C. (1993). *Statistics*. Dubuque, IA: Wm. C. Brown Publisher.