

Connecting IM Pattern and Selective Perceived Responsiveness to Relationship: A Cluster-Based Approach

Hao-Ping Lee

dimension4.cs03@nctu.edu.tw
National Chiao Tung University
Hsinchu, Taiwan

Chih-Heng Lin

cowbon.cs03@g2.nctu.edu.tw
National Chiao Tung University
Hsinchu, Taiwan

Kuan-Yin Chen

kevchentw.cs03@g2.nctu.edu.tw
National Chiao Tung University
Hsinchu, Taiwan

Yung-Ju Chang

armuro@cs.nctu.edu.tw
National Chiao Tung University
Hsinchu, Taiwan

ABSTRACT

As people utilize instant messaging (IM) to communicate with people of various relationships, they pay different amounts of attention to and have different communication practices with them of different relationships. However, we haven't seen a close investigation of how users' IM communication patterns relate to different groups of IM contacts. We collected IM logs of 547 sender-recipient pairs from 33 smartphone users over the course of 4 weeks, and used k-mean clustering to identify 6 clusters of these users' IM communication patterns. We illustrate the characteristics of the IM patterns of these distinct clusters as well as how the patterns relate to the relationship between the senders and the recipients within these clusters respectively.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**.

KEYWORDS

responsiveness, instant messaging, k-means algorithm, clustering

ACM Reference Format:

Hao-Ping Lee, Kuan-Yin Chen, Chih-Heng Lin, and Yung-Ju Chang. 2019. Connecting IM Pattern and Selective Perceived Responsiveness to Relationship: A Cluster-Based Approach. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct)*, September 9–13, 2019, London, United Kingdom. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3341162.3344841>

1 INTRODUCTION

People nowadays rely on instant messaging (IM) as a primary channel to communicate with people within their social circles. Prior research shows that smartphone users respond to incoming messages from different contacts differently [5] and they also perceived different levels of disruption [4, 5, 10] from different social contacts. Recently, Lee *et al.* [3] show that both responsiveness and perception of disruption are affected by the mutual expectations between senders and recipients. However, to date, we still have little understanding of the characteristics of the communication patterns between senders and recipients with whom they have different relationships. While prior works [1, 9] have attempted to predict tie strengths between senders and recipients using IM message history, different from them, we aim to use an unsupervised-machine-learning approach to identify distinct clusters of IM patterns. Our objective was to get insights into the important characteristics of the IM patterns between senders and recipients within these clusters, and examine how these patterns related to sender-recipient relationship and to recipient's selective perceived responsiveness to senders. Using the k-means clustering technique, we identified 6 clusters that characterized distinct IM patterns between senders and recipients. In addition, the results

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6869-8/19/09...\$15.00

<https://doi.org/10.1145/3341162.3344841>

Table 1: List of features related to IM communication pattern

<i>IM Communication Pattern</i>	
Intensity	{AVG,MIN,MAX,STD}
	# of msg per day
	MAX # msg in an hour
	MAX # msg in an hour / AVG # msg per hour
	# of incoming msg / # of total msg
Regularity	Average msg length
	Ratio of days have msg
	{AVG,MIN,MAX,STD}
	# hours have msg per day

show a connection between IM patterns and both senders-recipient relationship and perceived selective responsiveness, respectively.

2 DATA COLLECTION

Detail of the study procedure and participants have been reported in [3]. Here we briefly summarize the key parts. We analyzed IM history log data from 33 (17 males and 16 females) of 34 participants, including 20 students and 13 non-students, aging 20 to 50 ($M=25$) as one participant did not provide any of her IM logs. All of them were active users of Facebook and/or Line Messenger, the two most popular IM apps in Taiwan. These study participants participated in our field study for at least four weeks ($M=35.6$ days), in which they reported their relationship information and perceived responsiveness with 20 IM contacts using an online questionnaire. Among these 20 IM contacts, 10 of them were selected by participants with the instruction of those they expected to communicate with in the upcoming study period. An extra 10 IM contacts were selected by the research team later in the study based on participants' sampled IM contacts during the study. After the field study, they also provided us with their IM logs with 20 chosen IM contacts in text files. For each IM log, we extracted IM communication patterns using the method proposed in [1] to preserve the privacy of our participants. All of them received compensation of NT\$1,600 (approximately US\$52) for participating in the study, and an additional NT\$400 if they also participated in the post-study interview.

3 MEASURING IM PAIRS' PERCEIVED RESPONSIVENESS AND RELATIONSHIP

We aimed to investigate participants' **responsiveness**, i.e., how fast participants perceived themselves to be responding to the messages from their contacts, as it indicates their

selective attention to the incoming interruptions from those contacts. Furthermore, we included two facets of relationship in the contact questionnaire: **closeness** and **relationship type**, as the two could help us understand the composition of contacts in each cluster.

Responsiveness

We let participants report their perceived responsiveness to specific contacts using one of these options: *"I usually respond immediately/ in a couple of minutes/ in half an hour/ within an hour/ within a day/ don't respond within a day/ don't respond"*.

Closeness

Following Lee *et al.* [3], we adopted a 5-point Likert scale that asked *"How close do you feel to this person?"* to measure the closeness between participants and each of their chosen contacts.

Relationship Type

We divided the relationship type between participants and their contacts into Immediate family member (child, sibling, parent); Extended family member; Superior at work; Subordinate at work; Colleague; Client; Service provider; Friend; Acquaintance; Significant other; and Other. We grouped these relationship type into: Significant Other (SO), Social, Work and Family [2, 7, 8].

4 EXTRACTING IM PATTERN FEATURES

We only considered IM logs of which the participants and the selected contacts had exchanged IM messages during the study period. In total, we extracted IM patterns from the IM logs of 547 sender-recipient pairs. For each IM log, we extracted communication features following prior research [6, 9] that represented two IM communication patterns between the sender and the recipient: communication intensity and regularity, shown in Table 1.

5 CLUSTERING IM COMMUNICATION PATTERN

We used the 16 IM pattern features in Table 1 to form IM pattern groups using the k-means clustering algorithm¹. We rescaled features using min-max normalization² to avoid features with large differences from getting over-weighted in clustering. To find the best value of k in forming clusters, we iterated $k = [2:16]$, and evaluated each cluster configurations using within cluster sum of errors (WCSS) and Silhouette coefficient³. These two indicate the variation within clusters and between clusters, respectively. Figure 1 and Figure 2 together suggest that six was an optimal value for k, based

¹https://en.wikipedia.org/wiki/K-means_clustering

²https://en.wikipedia.org/wiki/Feature_scaling

³[https://en.wikipedia.org/wiki/Silhouette_\(clustering\)](https://en.wikipedia.org/wiki/Silhouette_(clustering))

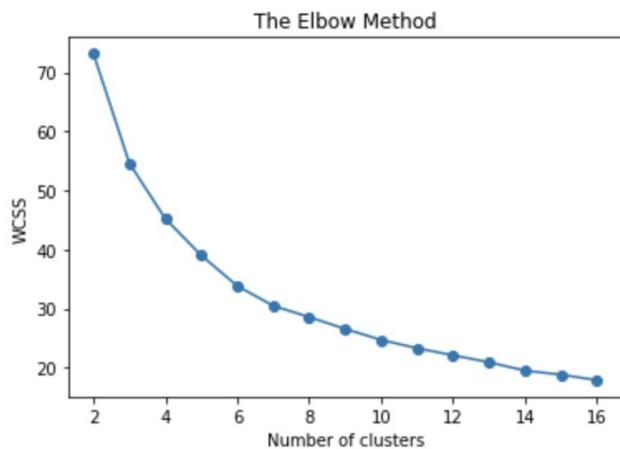


Figure 1: Result of WCSS of different k values, where we could observe an "elbow" when $k=6$.

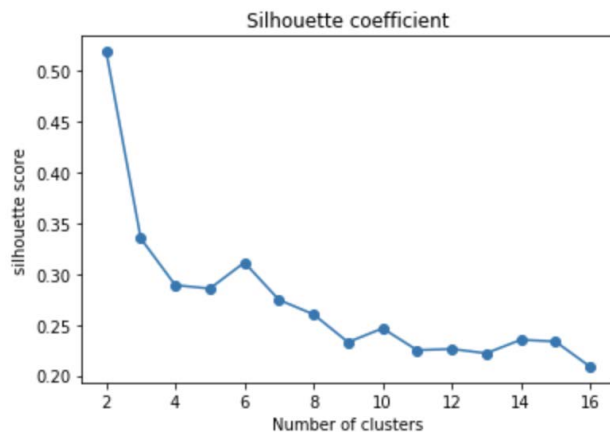


Figure 2: Result of Silhouette coefficient of different k values, where we could find a local maximum when $k=6$.

on the Elbow method⁴, and due to the fact that $k=6$ has the maximum Silhouette coefficient value when it is greater than 3.

6 IDENTIFYING 6 IM PATTERNS CLUSTERS AND THEIR RELATIONSHIP CHARACTERISTICS

Among all the sender-recipient pairs, 29.8% selected responding immediately and 28.15% selected responding within several minutes, respectively. However, the results, as shown in Table 2, show that the resulting six clusters display not only different IM communication patterns, but also different relationship characteristics and perceived responsiveness among the clusters.

⁴[https://en.wikipedia.org/wiki/Elbow_method_\(clustering\)](https://en.wikipedia.org/wiki/Elbow_method_(clustering))

Cluster 1: Active Contacts (N=130)

In this cluster, participants communicated actively with their contacts. They exchanged on an average of 9 messages daily with each other, during approximately half of the days (46.84%). The numbers of incoming and outgoing messages between them were generally balanced. Compared to other clusters, this cluster has the highest percentage of Family (13.85%) and Social (53.85%) among the relationship types, and the lowest percentage of SO (0.77%). It appears that this cluster describes an IM pattern of maintaining relationships via active message exchange. Compared to the overall perceived responsiveness, participants were less responsive to their contacts (28.46% responded immediately, 23.85% responded within several minutes).

Cluster 2: Distant Contacts (N=30)

Participants communicated with the contacts with the lowest intensity among all clusters. They exchanged less than a message daily (0.87), and only exchanged in 11.25% of the day. However, the messages being exchanged were of the longest length on average (40.23 words per message). This implies that while participants did not frequently communicate with the contacts using IM, each message had the most amount of content. The Family and Social relationship type appeared more often in this cluster. However, these contacts on average were more distant from the participants, with the lowest closeness of 2.53 (on a 5 Likert scale) on average among all clusters. Not surprisingly, participants also perceived themselves as the least responsive to these contacts (16.67% responded immediately, 26.67% responded within several minutes).

Cluster 3: Intimate Partners (N=9)

In this cluster, participants and the contacts exchanged messages with the highest intensity, on an average of 211 messages per day; with the highest regularity, almost every day (98.69% of days), and with the longest time span, 10 hours (10.28) per day. While they exchanged the largest number of messages, their messages were on average of the shortest length (6.86 words per message), suggesting frequent but brief message exchanges. Participants shared the highest closeness (5 on a 5 Likert scale) with the contacts in this cluster, indicating that these contacts were likely the participants' most intimate communication partners. Not surprisingly, SO appeared most often in this cluster (66.67%), followed by social (33.33%). Two-thirds of the participants in this cluster perceived themselves immediately responsive to these "intimate contacts" (66.67% responded immediately). Interestingly, none of the rest selected responding within several minutes, but responding within 30 minutes and an

Table 2: IM Communication Pattern, Relationship Characteristics and Responsiveness of Each Cluster

Clusters	Cluster 1 (N=130)	Cluster 2 (N=30)	Cluster 3 (N=9)	Cluster 4 (N=37)	Cluster 5 (N=97)	Cluster 6 (N=244)	All (N=547)
<i>AVG # of msg per day</i>	9.00	0.87	211.09	56.72	1.01	2.12	10.62
<i>Average msg length (characters)</i>	11.86	40.23	6.86	8.47	11.46	13.24	13.65
<i>Ratio of days have msg (%)</i>	46.84	11.25	98.69	86.06	5.81	15.99	27.36
<i>AVG # hours have msg per day</i>	1.10	0.13	10.28	4.74	0.09	0.30	0.91
<i>Closeness</i>	3.64	2.53	5.00	4.11	2.92	3.06	3.25
<i>SO (%)</i>	0.77	3.33	66.67	27.03	1.03	0.82	3.84
<i>Social (%)</i>	53.85	53.33	33.33	51.35	50.52	52.87	52.29
<i>Family (%)</i>	13.85	13.33	0	5.41	7.22	7.79	9.14
<i>Work (%)</i>	30.77	30.00	0	16.22	41.24	38.11	34.37
<i>Respond immediately (%)</i>	28.46	16.67	66.67	43.24	30.93	28.28	29.80
<i>Respond within couple minutes (%)</i>	23.85	26.67	0.00	29.73	30.93	30.33	28.15

hour. This suggests a relatively extreme self-perceived responding behavior: either responding immediately or slower than within several minutes.

Cluster 4: Close Buddies (N=37)

While participants shared the second highest IM communication intensity and regularity with the contacts, they exchanged much fewer messages with them compared to those to those "intimate partners" in the aforementioned cluster. They exchanged on average 56.72 daily messages. Participants and these contacts exchanged messages throughout of a day (86.06%), and their IM communication time spanned almost 5 hours (4.74) every day. Similar to participants and "intimate partners", they were in mostly SO and Social relationships, but in Social relationship type with much larger proportions (SO: 27.03%; Social: 51.35%). Participants also shared high closeness with them (4.11 on a 5 Likert scale), suggesting that they were participants' close chat partners though not the most intimate ones. Participants also quite often perceived themselves to be immediately responsive to these contacts (43.24%). 29.73% responded within several minutes.

Cluster 5: Functional Contacts (N=97)

In this cluster, participants exchanged messages with the contacts with relatively low intensity and regularity, on average 1.01 messages per day and merely 5.8% of the days, respectively. Interestingly, participants had the highest ratio of maximum messages exchanged within an hour to the average hourly messages with them. That is, participants and the contacts occasionally sent a burst of messages. This cluster had the highest percentage of Work (41.24%) than other clusters did. Within the cluster, the top category is Social (50.52%). The average closeness between participants

and the contacts was low (2.92 on a 5 Likert scale). We speculate that this cluster illustrates "functional conversations" that took place when participants and their contacts communicated for a specific purpose, thus aggregating a burst of messages when they happened. Participants perceived their responsiveness toward these "functional contacts" also to be higher than the average (30.93% respond immediately, 30.93% within several minutes) of all contacts.

Cluster 6: Passive Contacts (N=244)

In this cluster, the intensity and regularity of message exchanges between participants and the contacts were generally low, but higher than the "functional contacts" described earlier, with an average of 2.12 messages per day, 16% of days, respectively. Similar to the previous cluster, the cluster included relatively more Work contacts (38.11%), but they were closer to the participants (3.06 on a 5 Likert scale) than the "functional contacts". It is noteworthy that this cluster contained the largest number of sender-recipient pairs, suggesting a more typical and common conversation practice among the 33 participants and their contacts than other clusters. The perceived responsiveness in this cluster was also similar to the overall perceived responsive (28.28% responded immediately, 30.33% responded within several minutes).

7 SUMMARY

In this paper, we use k-means clustering to distinguish 6 kinds of IM patterns between 547 sender-recipient pairs, based on the 4-week IM logs provided by the participants. The distinct characteristics among these clusters suggest that smartphone users had different styles of IM messaging behavior with different contacts, and some of these behavioral patterns occur much often in one kind of relationship than the others. Our findings could be leveraged to the design of future IM services. For example, as our findings suggest that

users have different receptivity to messages from the contacts in different IM pattern clusters, IM applications could offer customized notifications of IM messages if a certain cluster is detected. For example, IM services could provide users with customized features when identifying the sender-receiver relationship characteristics between the user and sender of the incoming message. We believe these preliminary findings provide another empirical evidence regarding the connection between mobile computer-mediated communication and sender-recipient relationship. In future work, we will look into more features to get a more holistic view of this connection.

REFERENCES

- [1] Daniel Avrahami and Scott E. Hudson. 2006. Communication Characteristics of Instant Messaging: Effects and Predictions of Interpersonal Relationships. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 505–514. <https://doi.org/10.1145/1180875.1180954>
- [2] Shelly D. Farnham and Elizabeth F. Churchill. 2011. Faceted Identity, Faceted Lives: Social and Technical Issues with Being Yourself Online. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. ACM, New York, NY, USA, 359–368. <https://doi.org/10.1145/1958824.1958880>
- [3] Hao-Ping Lee, Kuan-Yin Chen, Chih-Heng Lin, Chia-Yu Chen, Yu-Lin Chung, Yung-Ju Chang, and Chien-Ru Sun. 2019. Does Who Matter?: Studying the Impact of Relationship Characteristics on Receptivity to Mobile IM Messages. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 526, 12 pages. <https://doi.org/10.1145/3290605.3300756>
- [4] Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. 2015. Designing Content-driven Intelligent Notification Mechanisms for Mobile Applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 813–824. <https://doi.org/10.1145/2750858.2807544>
- [5] Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. 2016. My Phone and Me: Understanding People's Receptivity to Mobile Notifications. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1021–1032. <https://doi.org/10.1145/2858036.2858566>
- [6] Jun-Ki Min, Jason Wiese, Jason I. Hong, and John Zimmerman. 2013. Mining Smartphone Data to Classify Life-facets of Social Relationships. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13)*. ACM, New York, NY, USA, 285–294. <https://doi.org/10.1145/2441776.2441810>
- [7] Fatih Kursat Ozenc and Shelly D. Farnham. 2011. Life "Modes" in Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 561–570. <https://doi.org/10.1145/1978942.1979022>
- [8] Jonathan Pettigrew. 2009. Text Messaging and Connectedness Within Close Interpersonal Relationships. *Marriage & Family Review* 45, 6-8 (2009), 697–716. <https://doi.org/10.1080/01494920903224269> arXiv:<https://doi.org/10.1080/01494920903224269>
- [10] Fengpeng Yuan, Xianyi Gao, and Janne Lindqvist. 2017. How Busy Are You?: Predicting the Interruptibility Intensity of Mobile Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5346–5360. <https://doi.org/10.1145/3025453.3025946>