Communication Characteristics of Instant Messaging: Effects and Predictions of Interpersonal Relationships

Daniel Avrahami and Scott E. Hudson
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh, PA 15213, USA
{ nx6, scott.hudson }@cs.cmu.edu

ABSTRACT
Instant Messaging is a popular medium for both social and work-related communication. In this paper we report an investigation of the effect of interpersonal relationship on underlying basic communication characteristics (such as messaging rate and duration) using a large corpus of instant messages. Our results show that communication characteristics differ significantly for communications between users who are in a work relationship and between users who are in a social relationship. We used our findings to inform the creation of statistical models that predict the relationship between users without the use of message content – achieving an accuracy of nearly 80% for one such model. We discuss the results of our analyses and potential uses of these models.

Categories and Subject Descriptors
H5.2 [**Information interfaces and presentation**]: User Interfaces; H1.2 [**Models and Principles**]: User/Machine Systems.

General Terms
Measurement, Performance, Experimentation, Human Factors

Keywords
Instant Messaging, IM, Communication Patterns, Interpersonal Relationships, Predictive Models.

1. INTRODUCTION
Over the past few years, the use of Instant Messaging, or *IM*, has been growing rapidly. IM programs, or clients, were created to facilitate one-on-one communication between a user and their list of contacts, commonly referred to as a 'buddy-list', by allowing them to easily send and receive short textual messages (*instant messages*). A recent report estimated that 12 billion instant messages are sent each day. Of those, nearly one billion messages are exchanged by 28 million business users [22]. As more and more people use IM for their social as well as their work-related communication, we wanted to investigate the effect of interpersonal relationship on basic characteristics of IM communication (such as duration of session, length of messages, and the rate at which messages are exchanged), independent of message content.

In this paper, we report the collection and analysis of a corpus containing over 90,000 real instant messages exchanged between 16 participants and over 400 of their buddies. Our findings show, for example, that users tend to communicate longer with their social buddies than with work buddies, but do so at a significantly slower pace. We then report on the use of our findings to inform the creation of two statistical models that predict the relationship between a user and their buddy based solely on basic communication characteristics. One of these models is able to predict, with accuracy of nearly 80%, whether a user and a buddy are in a work or social relationship. We conclude by discussing the practical implications of our findings and predictive models.

1.1 Background
In [8], Duck et al. describe the effect of interpersonal relationships on everyday communication. Using diary reports, they collected accounts of everyday spoken communication (either face-to-face or telephone) from over 1,700 students. Their analyses showed that interpersonal relationship type had significant effects on different aspects of communication, including the quality, purpose and perceived value of the communication. Feldstein describes the importance of cues such as tempo, pauses, speech rates and the frequency of turns, to the way in which participants in a conversation perceive each other [9].

The growing popularity of electronic communication, such as email, IM, and SMS (Short Message System), raises similar interesting questions as to whether different relationship types would result in differences in electronic communication.

In its early days, IM gained its widest use supporting social communication. Grinter and Palen, for example, reported that teenagers used IM primarily for socializing and planning social events [10]. As a result, when IM was introduced into the workplace, it was often met with resistance, being perceived as a medium suitable primarily for social communication [12], [19]. However, research showed that IM communication in the workplace has many uses and benefits complementing other communication mediums. These uses range from quick questions and clarifications, coordination and scheduling, to discussions of complex work (see [3],[11],[12],[14],[17]).

Figure 1 presents a single real IM session from our data, exchanged between one of our participants and one of their buddies, a co-worker. This session illustrates the lightweight nature of IM communication. In fewer than two minutes, and using no more than 12 messages, both participant and buddy were able to exchange brief greetings (messages# 1 and 3), coordinate a simple task (messages# 2,4,6,7), and apologize (message# 11) for a typing error made more than 30 seconds earlier (message# 8). This session also illustrates the use of abbreviations, loose grammar and minimal punctuation, prevalent in IM [17],[20].

IM is often described as a "near-synchronous" communication medium, placing it between synchronous communication mediums, such as speech, and asynchronous communication mediums, such as email. Voida et al. attribute a number of interesting behaviors of IM users, such as their need to acknowledge typing errors, to the tension between the near-synchronous yet still asynchronous and persistent nature of IM dialog [20]. Previous research has shown significant differences in IM communication resulting from the frequency of use (by an individual), as well as from the frequency of communication between a pair of users [14] (related differences were observed by [21] in face-to-face communication). We propose, however, that investigation is needed of the effect of *relationship* on IM communication. While interpersonal relationship might affect the use of grammar, abbreviations, or even the need to apologize for typos, in this work we wanted to examine its effect on more basic characteristics of IM by answering the following two research questions:

- What, if any, are the effects of interpersonal relationship on basic characteristics of IM communication? And,
- If such effects exist, can basic communication characteristics be used to predict the interpersonal relationship between a user and their buddy?

We will now start by describing the data collection method we used in order to answer those questions.

2. METHOD

2.1 Data Collection

The data collection process used in this work has been reported in [2]. However, for completeness, we briefly describe it here.

Our data were collected using a custom plug-in for Trillian Pro, a commercial IM client developed by Cerulean Studios [6], which runs on Windows operating system. We used Trillian Pro as it supports the development of dedicated plug-ins through a Software Development Kit (SDK). Trillian Pro allows a user to connect to any of the major IM services (ICQ, AOL, MSN, Yahoo!, and IRC) as well as other services such as Jabber and Lotus Sametime [16] from within one application, thus allowing us to recruit participants without concern for the specific IM service they are using. (In fact, 8 of the 16 participants in our study used Trillian to communicate with buddies over two or more IM services during their participation.)

We decided to use a commercial client rather than develop a client on our own as it provides functionality beyond the simple exchange of text messages. For example, it allows file sharing, audio and video chats, sending images, etc. This reduced the likelihood of participants using other IM clients during the course of their participation in our study.

To capture instant messaging events, a copy of Trillian Pro was purchased for each participant and the data recording plug-in was installed. Our plug-in is written in C and implemented as a Dynamically-Linked-Library (DLL) that is run from inside Trillian Pro. The plug-in automatically starts and stops whenever Trillian Pro is started or stopped by the participant.

#	Time		Message Text
1	17:42:45	B:	Hey [*Participant's name*]
2	17:42:56	B:	what time does your group get in the AM?
3	17:42:57	P:	hey
4	17:43:01	P:	usually around 10
5	17:43:25	B:	ok
6	17:43:38	B:	i want to start circulating the card in the AM
7	17:43:58	P:	ok, good idea
8*	17:44:02	P:	that's for coordinating this
9	17:44:13	B:	no problem
10	17:44:27	P:	thanks :-)
11	17:44:35	P:	sorry bout the typo
12	17:44:38	B:	is ok

* *The participant meant to write "thanks" and not "that's"*

Figure 1. A single IM session between one of our participants (P) and a buddy who is their co-worker (B).

The following IM events are recorded:

- Message sent or received
- Trillian start or stop
- Message window open or close
- Starting to type a message
- Status changes (online, away, occupied, etc.) of both participants' and buddies'
- Incoming message indicator is blinking (if this setting is used)

Other events, described in [2], were also captured but are not relevant for the work presented here. All events were saved into log files along with the time in which they occurred. These log files were compressed "on-the-fly" by the plug-in, encrypted, and stored locally on participants' machines.

Participants were instructed to use Trillian Pro for all their IM interactions for a period of at least four weeks. Towards the end of their participation, each participant used a small coding program to indicate their relationship with each buddy in their buddy-list using the following 12 possible relationships: Co-worker (senior), Co-worker (peer), Co-worker (junior), Co-worker (other), Friend & Co-worker, Acquaintance, Friend, Family, Significant-other, Spouse, Self, and Bot. (A Bot is a computer program that users can communicate with through IM.) The compressed log files, along with the coding, were collected from participants' computers at the end of their participation and instructions were given to them for removing the plug-in.

Privacy of Data

A number of measures have been taken to preserve, as much as possible, the privacy of participants and their buddies. The text of messages was not recorded unless we received specific permission from the participants. Otherwise, messages were masked in the following fashion: Each alpha character was substituted with the character 'A' and every digit was substituted with the character 'D'. Punctuation was left intact. For example, the message "my PIN is 1234 :-)" was recorded as "AA AAA AA DDDD :-)".

When a participant opened a message window to a buddy for the first time (and that buddy was online), an alert was sent to the buddy notifying them of the participation in the study. Buddies of participants who had provided the additional permission to record the text of messages were notified with a different alert message that instructed them of a simple mechanism that allowed them to temporarily mask messages.

Finally, for determining that two events were associated with the same buddy, we created a unique ID for each buddy (using an MD5 cryptographic hash) and stored the ID of the buddy instead of the buddy-name itself.

2.2 Measures

2.2.1 *Relationship Categories*

Since our main interest was in interpersonal communication, the relationships classified as Self and Bot were excluded from further analysis. The remaining ten relationships, assigned by the participants, were grouped into the following three higher-level relationship categories: Co-worker (senior), Co-worker (peer), Co-worker (junior), and Co-worker (other) were categorized as *Work*. Friend, Family, Significant-Other, and Spouse were categorized as *Social*. Friend & Co-worker was categorized as *Mix* and so was Acquaintance.

2.2.2 *Defining IM Sessions*

We define an IM session to be a set of instant messages that are exchanged within a certain time proximity of one another. Unlike a conversation, a session is not determined by the content of its messages. Indeed, a single conversation may extend over multiple sessions, while a particular session may contain many conversations. Following Isaacs et al. [14], we categorized two instant messages as belonging to the same IM session if they were exchanged between a participant and their buddy within 5 minutes of one another.

2.2.3 *Communication Measures*

For each IM session, we computed a set of 12 measures describing basic characteristics of the session. They are:

- **Duration**: The length of time between the first and last message in the session (in minutes).
- **Message count**: The total number of messages exchanged in the session.
- **Turn count**: The total number of turns taken in the session. A single turn consists of consecutive messages sent by the same user.
- **Character count**: The total number of characters exchanged in the session (including spaces).
- **Messages-per-Minute**: The average number of messages sent per minute (Message count divided by Duration).
- **Messages-per-Turn**: The average number of messages sent per turn (Message count divided by Turn count).
- **Characters-per-Message**: The average length of messages (Character count divided by Message count).

Table 1. Session variables computed for the session presented in Figure 1.

Variable	Value	
Group	Student	
Relationship	Work	
Duration	1.88	minutes
Message Count	12	
Turn Count	7	
Character Count	232	
Messages per Minute	6.4	
Messages per Turn	1.71	
Characters per Message	19.3	
Seconds Until First Reply	1	seconds
Minimum Gap (between turns)	1	seconds
Maximum Gap (between turns)	24	seconds
Average Gap (between turns)	12.2	seconds
Time of Day	5:44	pm

- **Seconds Until First Reply**: The time between the end of the first turn and the beginning of the second turn (in seconds)*.
- **Minimum Gap**: The shortest gap between turns in the session (in seconds)*.
- **Maximum Gap**: The longest gap between turns in the session (in seconds)*.
- **Average Gap**: The average gap between turns in the session (in seconds)*.
- **Time of Day**: The time of the last message in the session.

To illustrate how these measures are computed, Table 1 shows the values of each of these measures computed for the transcript presented in Figure 1. For example, in this particular session the gap of 24 seconds between messages 4 and 5 represents the Maximum Gap. The ratio of Messages-per-Turn is 12 / 7 = 1.71, and the average message length (Characters-per-Message) is 232 / 12 = 19.3.

2.3 Participants

Data were recorded from 16 participants between May 2005 and September 2005. The participants included eight Masters students at our department and eight employees of a large industrial research laboratory, who used IM in the course of their everyday work. Of the latter group, six were full time employees (three first-line managers and three full-time researchers) and two were summer interns. We will refer to the first eight participants as the Students group, the six full-time employees as Researchers, and finally the two interns as Interns.

Of the Students, six were female and two male, with an average age of 24.5 (SD=2.39, Min=22, Max=29). Six of these

* The value of this variable cannot exceed 5 minutes, since a gap longer than 5 minutes would qualify as the end of the session.

Table 2. Overview of the data collected from each participation group

Participation Group	N	Avg age	Total hours recorded*	Avg hours recorded per participant per day	Total active buddies	Avg active buddies per participant	Total sessions	Sessions analyzed	Total msgs	Avg msg per recorded hour
Researchers	6	40.3	982.5	6.4	130	21.7	845	605	7290	7.4
Interns	2	34.5	373.0	5.6	61	30.5	757	543	10343	27.7
Students	8	24.5	3839.8	9.4	244	30.5	2903	2149	73906	19.2
Overall	16	31.7	5195.2	8.2	435	27.2	4505	3297	91539	17.6

* *Due to a number of corrupt log files, these numbers are slightly lower than the true value.*

participants ran the recording software on their personal laptops. One participant, who used a laptop at school and a desktop computer at home, ran the recording software on both machines. The eighth participant ran the recording software on his account on a shared desktop computer in the Masters students' lab. During their participation, each of these participants was engaged in a number of group projects as part of their studies.

The average age of the six Researchers was 40.33 (SD=4.97, Min=34, Max=49) with three female and three male. One female and one male, the average age of the Interns group was 34.5 (SD=3.54, Min=32, Max=37). The Researchers and Interns ran the recording software on their work laptops. For confidentiality reasons, we did not record the text of messages from any of the participants in the Researchers or Interns groups.

All of our participants except one were new to Trillian Pro but were able to automatically import the list of all their buddies into Trillian Pro. None of the participants had any difficulty making the transition to using Trillian Pro, although some assistance was required with customization of specific options to match the preferences that individual users were accustomed to. All participants ran the recording software for a period of at least 4 weeks. Two of the participants voluntarily continued their participation for a total of approximately 3 months.

2.4 Data Overview

Table 2 provides a summary of data collected. Using Trillian Pro as our data collection platform resulted in successful recording of a very high volume of IM events. (A small number of data files were unusable due to corruption in the on-the-fly compression, often as a result of participants' laptops running out of power.)

We collected a total of approximately 5200 hours of recorded data, observing over 90,000 incoming and outgoing instant messages assigned to over 4500 IM sessions between the participants and more than 400 buddies. Sessions ranged in duration from 2 seconds to 2.2 hours, and contained anywhere from 2 to 1098 messages. Two of the participants in the Researchers group recorded significantly fewer messages in their logs (96 and 350 messages). However, we did not remove their data from our models and analyses.

When their IM client was running, participants in the Students and Interns groups exchanged, on average, a single message every 2.2 and 3.1 minutes respectively. By comparison, the Researchers exchanged, on average, a single message every 8.1 minutes. Differences between the overall rates of message-exchanges by group were significant ($F[2,13]=5.08$, $p=.024$). A pair-wise comparison shows that the difference in the rate of messaging was significantly different between the Researchers and either the Students ($t(13)=-2.57$, $p=.023$) or Interns ($t(13)=2.71$, $p=.018$). There was no significant difference between the Interns and the Students groups ($p=.32$, N.S.).

2.4.1 Excluding Single-Turn Sessions

Single-turn sessions are IM sessions in which one user sends one or more messages without a reply. 1190 of the total sessions in our data were identified as single-turn sessions. (A large number

		Researcher s	Interns	Students
Work	Co-worker (senior)	22	6	1
	Co-worker (peer)	43	6	24
	Co-worker (junior)	34	-	2
	Co-worker (other)	9	-	-
Mix	Friend & Co-worker	16	13	80
	Acquaintance	-	2	12
Social	Friend	4	22	98
	Family	1	5	20
	Significant-other	-	3	2
	Spouse	-	2	-
Other	Self	1	1	5
	Bot	-	1	-

Table 3. Distribution of Buddies by Relationship and Group. (Note: A buddy appearing several times in a participant's buddy-list will also appear those many times in the data)

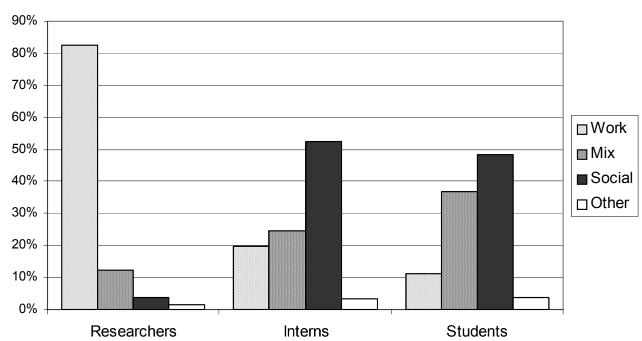

Figure 2. Distribution of Buddies by Relationship Category and Group.

of those represent failed communication attempts.) Since single-turn sessions provide very little information about the interaction between a participant and a buddy, we removed these sessions from all the analyses and modeling presented next. After excluding the single-turn sessions, our data set contained a total of 3297 sessions between 412 participant-buddy pairs.

2.4.2 Relationship Distribution

The distribution of relationships as indicated by our participants is presented in Table 3. We can see that some relationships appeared very little or were not reported at all by different participation groups. For example, our Researchers indicated 22 of their buddies as being in the Co-worker (senior) category, while only one buddy was identified in that category from the Students group. Figure 2 shows the proportion of each high-level relationship category as indicated by each participation group. (Note that if a buddy appears on a participant's buddy-list more than once using different buddy-names, then that buddy will also be counted more than once in the data.)

From both Table 3 and Figure 2, it is clear that the distribution of relationships is very different between our participation groups. For example, 83% of the buddies that our Researchers communicated with were identified as Work, compared to 11% for our Students group. Similarly, over 49% of the buddies in the Students and Interns groups were identified as Social, compared to only 4% for the Researchers. These differences between the participation groups were controlled for in the analysis.

3. RESULTS

Table 4 shows the correlation coefficients for each pair of measures. As could be expected, the correlation between Duration, Message count, Turn count, and Character count is extremely high ($r \geq .88$). It is also interesting to note that the inverse correlation between Messages-per-Minute and Average Gap is only $r=-.25$. The two are inversely correlated since, when message rate is higher, the gap between turns is likely to be shorter (recall, however, that message rate is related not only to gaps between turns, but also to gaps within turns).

To examine the effect of relationship on each of the communication characteristics variables described above, we used a repeated-measures analysis of variance (ANOVA) in which Relationship Category (Work, Mix, Social) and Group (Researchers, Interns, Students) were repeated. Because participants and buddies typically communicated with one another more than once, observations were not independent of one another. Participants and BuddyID were modeled as random effects. Further, since each participant belonged to only one participation group, Participants were nested in Group. Similarly, since buddies appeared, for the most part, on only a single participant's buddy list, BuddyID was nested first in Participants, then in Group. This analysis allowed us to control for differences in communication characteristics that originate from the differences between our participation groups (evident in Tables 2 and 3) or that originate from individual (or dyadic) differences.

Our results, summarized in Table 5, show that many of the communication characteristics were affected by the Relationship between the users and their buddies. Sessions of buddies in a Work relationship were shorter in duration – due in part to a smaller number of messages exchanged and to an overall faster exchange, although the length of messages themselves was longer. Here are the results in detail.

We found that Relationship had significant effect on Duration ($F\ [2,331] = 8.04$, $p<.001$). Sessions between buddies in a Social relationship lasted, on average, 2 and a half minutes longer (M=6.6 minutes) than sessions between buddies in a Work relationship (M=4 minutes) and about one and a half minutes longer than sessions between buddies in a Mix relationship (M=5.2 minutes)[1]. A planned pair-wise comparison showed that Duration of session was significantly different between sessions with buddies in a Social relationship and sessions with buddies in either Work or Mix relationships[2] ($t(310)=3.65$, $p<.001$, and $t(331)=2.72$, $p=.007$, respectively). Since Duration, Message count, Turn count, and Character count are all correlated at over .85 (see Table 4) one could expect similar differences for these variables too. This is indeed true for the pair-wise comparisons between Social and Work relationships (Message count M=25.9 vs. M=13.8; $t(382)=3.27$, $p=.001$; Turn count M=15.3 vs. M=8.8; $t(350)=3.28$, $p=.001$; and Character count M=844.6 vs. M=459.5; $t(316)=2.95$, $p<.004$) but not for the Mix relationship.

We found that Relationship had significant effect on Messages-per-Minute ($F\ [2,99] = 4.75$, $p=.01$). Interestingly, we discovered that while users tended to have longer sessions with buddies in a Social relationship and exchanged more messages per session, they exchanged messages with these buddies at a significantly slower pace. Messages-per-Minute was significantly lower for buddies in a Social relationship compared to Mix relationship (M=4.6 vs. M=6.2 messages per minute; $t(115)=-2.99$, $p=.003$) and marginally significant compared to Work relationships (M=4.6 vs. M=6.0 messages per minute; $t(70)=-1.8$, $p=.078$). Messages-per-Minute did not vary significantly between Work and Mix.

A potentially related result is the significant effect of Relationship on Maximum Gap ($F\ [2,173] = 3.25$. $p<.05$), where a significantly longer maximum gap between turns was "allowed" in sessions with Social buddies (M=82 seconds) compared to sessions with Work buddies (M=69 seconds; $t(172)=-2.51$, $p=.013$). It is possible that the difference in Maximum Gap simply results from the fact that longer gaps are more likely in longer sessions that contain more turns. The correlation of $r=.46$ between Maximum Gap and the overall Duration of the session suggests that this explanation can account for a large portion of this effect but might not account for it entirely.

Our results also show that Relationship had a significant effect on Characters-per-Message ($F[2,229] = 7.85$, $p<.001$). The length of messages exchanged between buddies in a Work relationship were longer, on average, than messages exchanged between buddies in either a Mix or a Social relationship (M=38 vs. M=32 or M=30; $t(1,219)=3.95$, $p<.001$ and $t(1,250)=3.11$, $p=.002$).

[1] Because the independent variables were not completely orthogonal, we used Least Squared Means (LS Means) to control for the values of the other independent variables. The means reported throughout this article are LS Means.

[2] All pair-wise comparisons were done using the Tukey HSD post-hoc test.

Table 4. Correlation coefficients of the IM characteristics variables (N=3297)

	Duration	Message Count	Turn Count	Char Count	Message per Minute	Message per Turn	Chars per Message	Secs Until First Reply	Minimum Gap	Maximum Gap	Average Gap
Message Count	0.88										
Turn count	0.88	0.99									
Character count	0.88	0.95	0.95								
Message per Minute	-0.15	-0.04	-0.04	-0.05							
Message per Turn	0.15	0.16	0.07	0.12	-0.06						
Chars per Message	0.11	0.03	0.05	0.18	0.01	-0.06					
Seconds Until First Reply	0.03	-0.08	-0.09	-0.08	-0.20	0.09	0.04				
Minimum Gap	-0.12	-0.16	-0.17	-0.15	-0.12	0.07	0.05	0.56			
Maximum Gap	0.46	0.22	0.22	0.22	-0.30	0.12	0.07	0.49	0.25		
Average Gap	-0.01	-0.17	-0.18	-0.14	-0.25	0.09	0.08	0.69	0.87	0.58	
Time of Day	0.17	0.17	0.16	0.12	0.01	0.16	-0.11	0.00	0.00	0.08	-0.01

Table 5. The effect of relationship on IM characteristics (N=3297)

	Relationship Category						Analysis of Variance		
	Work		Mix		Social				
Variables	Mean	StdErr	Mean	StdErr	Mean	StdErr	*F*	d.f.	*p*
Duration (in minutes)	4.0	0.6	5.2	0.6	6.6	0.5	8.04	2/331	<.001
Message count	13.8	3.0	19.8	3.1	25.9	2.8	6.11	2/398	<.01
Turn count	8.8	1.7	12.2	1.7	15.3	1.6	5.96	2/374	<.01
Character count	459.5	122.7	673.6	123.6	844.6	115.2	4.71	2/340	<.01
Messages-per-Minute	6.0	0.5	6.2	0.4	4.6	0.4	4.75	2/99	<.05
Messages-per-Turn §	1.5	0.05	1.5	0.05	1.6	0.05	2.32	2/312	
Characters-per-Message	37.9	2.5	31.5	2.5	30.1	2.4	7.85	2/229	<.001
Seconds Until First Reply	36.9	3.0	35.0	3.1	36.0	2.7	0.11	2/151	
Minimum Gap (between turns)	12.0	1.8	12.4	1.9	12.1	1.6	0.02	2/111	
Maximum Gap (between turns)	68.7	3.8	77.0	3.9	81.8	3.4	3.25	2/173	<.05
Average Gap (between turns)	28.8	2.2	28.3	2.3	29.2	2.0	0.10	2/181	
Time of Day §	14.6	0.4	14.6	0.4	14.7	0.4	0.04	2/253	

§ - Participation Group (Researchers, Interns, and Students) had significant effect on this variable

Message length did not vary significantly between Mix and Social relationships.

We did not find significant effects of Relationship on any of the remaining communication characteristics variables. We did, however, find two significant effects of Participation Group on communication characteristics.

Participation Group had a significant effect on the average number of messages per turn (Messages-per-Turn) ($F [2,16] = 7.82$, $p<.01$), with the Students exchanging significantly more messages per turn than the Researchers ($M=1.7$ vs. $M=1.4$; $t(22)=-3.63$, $p<.002$). Messages-per-Turn was not significantly different between Interns and Students nor Interns and Researchers. This result is similar to results reported by Isaacs et al. where message exchange rate between their Light and Heavy IM users differed significantly [14] (in their work, they used the term “turn” to refer to what we consider a single message). Paraphrasing their terminology, underlying differences between our participation groups, and in particular the Researchers and Students, could warrant classifying them as Heavy and Super-Heavy respectively (see Table 2).

Participation Group also had significant effect on Time of Day ($F [2,15] = 36.8$, $p<.001$). This is not surprising considering that unlike the Students, the Researchers and Interns used IM primarily during business hours. This result is in accordance with results found by Begole et al. [5].

3.1 Discussion

Our analysis showed the significant effect of relationship on a number of the communication characteristics we investigated. We were not surprised by the effect of relationship on the overall length of sessions (including duration, number of messages, turns, and characters). However, we were surprised and intrigued by the effect of relationship on message exchange rate (Messages-per-Minute).

Based on our findings, the difference in message exchange rate between Work and Social relationships cannot simply be accounted for by differences in the length of messages. In fact, our results show the exact opposite. Not only did participants and buddies in a Work relationship exchange longer messages on average, but they also did so at a faster pace overall. An interesting possible explanation for the differences in pace is that users devoted different levels of their attention to the different conversations. In other words, it is possible that users focus less of their undivided attention to conversations with their Social buddies and give more attention to conversations with their Work buddies. This explanation is supported, in part, by the significant differences in the Maximum Gap between turns. The Maximum Gap reflects the maximum time that users let their conversation partners wait before responding. The significantly higher gap allowed between buddies with a relationship of a social nature may again suggest that less focus of attention is given to sessions with those buddies in comparison to conversation with buddies in a work relationship.

One possible explanation for the interesting differences in message length is that conversation between buddies in a work relationship is less casual and users construct their ideas more carefully before sending them. Another explanation could be that conversation with work buddies requires greater verbosity to achieve common ground than conversation with social buddies. Finally, it is possible that the concepts discussed with work relationships (perhaps more complex) simply require the use of longer terms to describe.

Having examined the effect of relationship on communication characteristics, we wanted to see whether the process could be reversed such that the differences in the communication characteristics are used by a classifier to predict the relationship between users and their buddies. Since the communication characteristics we examined do not use the content of messages, such a classifier would not pose too great an invasion of privacy. We now describe the creation of two such classifiers.

4. PREDICTING RELATIONSHIPS

In this section we describe the creation of two predictive models (or "classifiers") that predict the relationship between a user and their buddies using only those basic characteristics shown in the previous section. Both models were generated using Nominal Logistic Regression. (Other classification techniques, including Naïve Bayes and Decision Trees were also explored but resulted in lower accuracy.) Both models used a similar two-step process to provide their predictions. In the first step, the model predicts the relationship for each individual IM session, and in the second, a majority vote is taken for each participant-buddy relationship, across all their joint sessions, to provide a final classification.

It is important to stress that our models attempt to predict the *relationship* between IM buddies, not the content of their individual conversations (although the two are undoubtedly related). That is, a model should classify friends as being in a Social relationship even if they sometimes talk about work. Similarly, a model should classify co-workers as being in a Work relationship even though they may discuss the location for an after work drink.

Predictive models of the relationship between IM users can be used in a number of ways. First, predictions of relationship could be used to augment IM systems. For example, a system such as Lilsys [4] could set indicators of unavailability to buddies individually, based on predicted relationships. IM clients could also alert users to incoming messages differently, depending on their predicted relationship with the sender. An augmented IM client that observes the content of incoming messages (similar to QnA [1]), or a client that predicts whether a user is likely to respond to a message (using models such as the ones presented by Avrahami and Hudson in [2]) could use predictions of relationship to help guide whether or not to increase the salience of incoming messages. A completely different category of uses for these predictive models would be to allow their predictions, originating in IM, to propagate to other communication mediums. With many of today's IM service providers, such as Microsoft, AOL, Yahoo! and Google also providing email (and recently also Voice-over-IP), a person's IM identity (their buddy name) is often their email identity as well. Thus, a prediction of the relationship with a person, based on their IM interaction, could be used to enhance the interaction with the same person in different mediums. For example, such predictions could be used to inform systems such as the Priorities system that predict email interaction [13]. Finally, predictive models of relationships could also be used to provide an overview of IM communication in a whole organization, and even comparison between organizations.

We now describe this process in detail followed by results and prediction accuracy.

4.1 Preparing the Data

Informed by the results presented in the previous section, we used the following 8 variables (or *features*) in our predictive models: Duration, Message count, Turn count, Character count, Messages-per-Minute, Messages-per-Turn, Characters-per-Message, and Maximum Gap. We could not use (or control for) Group or Participant in the models as these are not independent of relationship. We felt that, in order for these models to be interesting, they must work well across groups and without knowledge of the group that a participant belongs to (otherwise, if one knows, for example, that a participant belongs to the Researchers group then one could simply guess that the relationship with a buddy is a Work relationship and be correct 84% of the time). In order to make up for the inability to control for differences between the groups and participants, we applied a natural-log transformation to each of our variables (except for variables that represent rates). Thus, our final set of variables was as follows: log(Duration), log(Message count), log(Turn count), log(Character count), Messages-per-Minute, Messages-per-Turn, Characters-per-Message, and log(Maximum Gap).

SN	Buddy	Actual	Session Variables	Prediction
i	*b1*	0 *(Work)*	$< v_1, v_2, \ldots, v_n >$	0 *(Work)*
i	*b1*	0 *(Work)*	$< v_1, v_2, \ldots, v_n >$	0 *(Work)*
i	*b1*	0 *(Work)*	$< v_1, v_2, \ldots, v_n >$	1 *(Social)*
i	*b2*	1 *(Social)*	$< v_1, v_2, \ldots, v_n >$	0 *(Work)*
i	*b2*	1 *(Social)*	$< v_1, v_2, \ldots, v_n >$	1 *(Social)*

(a) Step 1: Predict Relationship for each Session

SN	Buddy	*n*	Actual	Average	Final Prediction	Correct
i	*b1*	3	0 *(Work)*	.333	0 *(Work)*	Yes
i	*b2*	2	1 *(Social)*	.5	0 *(Work)*	No

(b) Step 2: Predict Relationship for each Buddy using average of individual Session predictions

Figure 3. Classification Process illustration: (a) Session-level predictions and (b) final Buddy-level predictions with one correct and one incorrect predictions.

4.2 Model 1: Work vs. Social

The first of the two models classifies relationships into one of two classes: Work or Social. For this model, we used a subset of our data containing only sessions between participants and buddies in either a Work or Social relationship. This subset contained 2379 sessions with 292 participant-buddy pairs (of which 203, or 70%, appeared in two sessions or more).

To test the accuracy of the model, we used a 16-fold cross validation method. That is, the model is created over 16 trials, one trial for each participant, and the combined accuracy is reported. Typically, with cross-validation, the data are randomly divided into a number of subsets. In our case, however, different sessions from the same participant are not independent (especially sessions with the same buddy), and randomly segmenting the data would likely result in some of a participants' sessions appearing in both the training and test data. This would give the model an unfair (and unrealistic) advantage. Instead, we used a more conservative cross-validation method in which, for each trial, the full data of a single participant is excluded as a test set and the data from the other participants are used for training.

4.2.1 Training Process

The training process for each trial follows three steps: First, all sessions of one participant are excluded and kept as a test set. Next, the remaining data are adjusted to contain an equal number of sessions for each class (described below). Finally, the model is generated using the sessions in the training set.

Adjusting the distribution of the training set is important in order to prevent the underlying bias in the distribution of sessions from biasing the predictions of relationships (for example, while only 37% of the buddies were identified by our participants as in a Social relationship, over 45% of sessions recorded were with those buddies). This bias in distribution was mostly a result of variance in the amount of data recorded from the different participants. Participants in the Researchers and Interns groups, for example, tended to use IM during business hours on weekdays, while participants in the Students group used IM nearly 24 hours a day, 7 days a week. As a result, our data contain a greater number of sessions from our Students. Thus, prior to training a model, the training set is adjusted to include an equal number of sessions for each relationship category. This prevents the model from merely classifying relationships as Social as a result of their high frequency in the data. For example, if the training set consists of 700 Work sessions and 800 Social sessions, then 100 Social sessions are selected at random and excluded from the training set.

	Classified as Work	Classified as Social
Work	40.9% (83)	5.9% (12)
Social	14.8% (30)	38.4% (78)
	Accuracy: 79.3%	

Figure 4. Classification results of a model predicting Work vs. Social relationships.

4.2.2 Classification Process

The classification used in our models follows a two-step process (illustrated in Figure 3). First, the model is used to provide a relationship prediction of 0 (Work) or 1 (Social) for each session in the test set (Figure 3a). We will refer to these predictions as "Session-level predictions". In the second step (Figure 3b), a single final prediction is provided for each buddy using a majority vote among all session-level predictions for the same buddy. In other words, the model provides a final prediction based on whether the average session-level prediction is greater or smaller than 0.5. The second step is performed only for buddies with whom a participant had two or more sessions. In case of a tie (the average equals 0.5), the majority prediction of all session-level predictions (for all buddies) is assigned as the final prediction for the buddy. Figure 3b includes an illustration of a case where a tie is resolved (in this case, to generate an incorrect classification).

4.2.3 Performance Results (Model 1)

The performance of this first model, for buddies with two or more sessions, is presented in Figure 4. The model was able to accurately predict 161 of the 203 relationships, for an accuracy of 79.3%; significantly better than the 53.2% prior probability (G^2 (1,203)=73, p<.001). (Prior probability represents the accuracy of a model that picks the most frequent answer at all times.)

We were curious to see the model's performance when classifying relationships for buddies with whom our participants communicated only once. As expected, the accuracy of these predictions was much lower (41.6%). We believe that it is not unreasonable, however, for a system using such a model to require at least two data points before providing a final prediction of relationship.

	Classified as		
	Work	Mix	Social
Work	25.3% (74)	5.1% (15)	2.0% (6)
Mix	8.2% (24)	14.7% (43)	7.8% (23)
Social	9.6% (28)	17.1% (50)	10.2% (30)
	Overall Accuracy: 50.2% Work vs. Rest: 75.1% Social vs. Rest: 63.5%		

Figure 5. Classification results of a model predicting Work vs. Mix vs. Social relationships.

4.3 Model 2: Work, Mix, Social

Since our full data set consisted also of buddies with whom our participants were in a relationship that was a mix of both social and work, we next attempted the much harder 3-way classification problem. For this model, we used the full data set, which contained 3297 sessions with 412 participant-buddy pairs (of which 293, or 71%, appeared in two or more sessions). Again, we used a 16-fold cross-validation, excluding the data from one participant each time, and training on the remaining data. The combined accuracy of the 16 trials is reported.

4.3.1 Training Process

The training process was almost identical to the process used for the 2-way model. In addition to adjusting the training set to contain an equal number of Work and Social sessions, training sets were adjusted to also include an equal number of Mix sessions.

4.3.2 Classification Process

Again, a two-step classification process is used, similar to the process described earlier. In the first step, the model provides a relationship prediction of 0 (Work), 0.5 (Mix), or 1 (Social) for each session in the test set. In the second step, a single final prediction is provided for each buddy using a slightly modified voting step among all session-level predictions for the same buddy.

4.3.3 Performance Results (Model 2)

The performance of this second model is presented in Figure 5. The model was able to accurately predict 147 of the 293 relationships. This model's accuracy was only 50.2% (compared to the prior probability of 36.9%). Again, the accuracy of predictions for buddies with whom our participants communicated only once was even lower (36.1%). A closer examination of the model's predictions shows that the model was much more accurate at distinguishing Work from not Work (75.1%) than it was at distinguishing Social from not Social (63.5%).

4.4 Discussion

The performance of our first model (predicting Work vs. Social) was surprisingly high considering that no content of messages was used to generate the predictions. The drop in accuracy when moving to the 3-way model (predicting Work vs. Mix vs. Social) could be a result of the greater difficulty of a 3-way classification in general. However, we believe that the main reason for this drop in accuracy is that the Mix relationship is, indeed, similar to both the Work and Social relationships. We are currently examining the possibility of using a cascading approach, in which a model first predicts whether a relationship is Work or not, then a second model attempts to distinguish Mix from Social.

Indeed it is possible that the features used by our models are simply insufficient for distinguishing between all three of the relationship categories. This may suggest that different features are needed in order to accurately distinguish between the three categories, and in particular distinguish Mix from Social. These features may need to use some aspects of the content of messages (for example, using the Linguistic Inquiry and Word Count program [18]). Still, these models present an exciting potential for predicting relationships without using the private and potentially sensitive content of messages.

5. SUMMARY & FUTURE WORK

In this paper we described an analysis of the effect of interpersonal relationship on basic characteristics of IM communication. We describe, for example, a number of results that suggest that, while IM sessions with social contacts are longer in duration, users focus, on average, less of their undivided attention to these sessions. Our findings add to previous research, which showed the effect of interpersonal relationships on face-to-face and phone communication, by extending it to IM communication. This work also complements previous research that described the effect of frequency of communication on basic characteristics of communication in both synchronous and asynchronous mediums.

We used the results of our analysis to inform the creation of two predictive models. One of the models described was able to predict, with 79.3% accuracy, whether a user and a buddy are in a work or social relationship. This accuracy is impressively high considering that only basic characteristics of communication were used, without knowledge of the actual content of messages. Finally, we discussed our results and potential uses for predictive models of interpersonal relationship.

Using a sample of 16 participants meant that our data set, while not small, contained conversations between only 412 participant-buddy pairs. We plan a new data collection phase for the near future in order to examine the application of the results presented here to a new set of participants (we are currently collecting data from 11 additional participants, including 4 employees of a local startup company). Still, we believe that our findings should generalize beyond the 412 pairs in our set. Specifically, the relatively high performance of our first predictive model, despite the significant differences between our participation groups (in age, profession, composition of buddy-list, etc.), suggests a robustness of our underlying findings.

In the work presented in this paper, we grouped the fine-grain relationship categories presented in Table 3 into three high-level categories (Work, Mix, and Social). This grouping was done, in part, due to the uneven distribution of fine-grain relationships in our data. In the next data collection phase, we plan to expand the list of relationships to also include types shown by previous literature as having distinct properties (such as Best Friend). We then plan to examine, in detail, the effect of fine grain relationship categories on communication (e.g., do communication characteristics differ between sessions with a peer and with a senior co-worker?). However, it is important to remember that, from a machine-learning perspective, attempting to classify closely related concepts can be very difficult. As the performance of our models dropped with the introduction of the Mix relationship, one can expect a classification of all 10 fine-grain relationships to be very difficult.

Kraut et al. showed that physical distance has significant effect on coordination and communication [15]. We are interested in examining whether and how physical distance between IM buddies affects their basic communication characteristics. We plan to use the scale from Cummings and Ghosh [7] to get a coding of distance from future participants. We suspect that interesting differences exist in the interaction of relationship and physical distance.

In conclusion, Instant Messaging is maturing and with it, its users. The young adults who have been using IM for their social communication for over a decade are now joining the workforce. Thus, a better understanding of the factors affecting IM communication is needed. More specifically, a better understanding of the differences and similarities between social and work IM and of how the two may coexist. We believe that the work described in this paper is an important step towards reaching this goal.

6. ACKNOWLEDGMENTS

We would like to thank Laura Dabbish, Sue Fussell, Darren Gergle, Eric Horvitz, and Bob Kraut, for their useful comments and suggestions. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010, and by the National Science Foundation under grants IIS 0121560 and IIS 0325351.

7. REFERENCES

[1] Avrahami D., and Hudson S. E. QnA: Augmenting an instant messaging client to balance user responsiveness and performance. *In Proceedings of CSCW ‘04*. ACM Press, NY, 2004, 515-518.

[2] Avrahami D., and Hudson S. E. Responsiveness in Instant Messaging: Predictive Models Supporting Inter-Personal Communication. *In Proceedings of CHI ‘06*. ACM Press, NY, (in press).

[3] Bradner, E., Kellogg, W.A., and Erickson, T. The adoption and use of 'Babble': A field study of chat in the workplace, *In Proceedings of ECSCW ‘99*. Springer Publishing, 1999, 139-158.

[4] Begole J., Matsakis N. E., and Tang J. C. Lilsys: Sensing unavailability. *In Proceedings of CSCW ‘04*. ACM Press, NY, 2004, 511-514.

[5] Begole, J., Tang, J.C., Smith, R.E., and Yankelovich, N. Work rhythms: Analyzing visualizations of awareness histories of distributed groups. *In Proceedings of CSCW ‘02*. ACM Press, NY, 2002, 334-343.

[6] Cerulean Studios http://www.trillian.cc

[7] Cummings, J., and Ghosh, T. Boundary-spanning ties for coordination at a distance: A relational model of communication media use in teams. (under review).

[8] Duck, S., Turr, D., Hurst, M., and Strejc, H. Some evident truths about conversations in everyday relationships: All communications are not created equal. *Human Communication Research 18*, 2 (1991), 228–267.

[9] Feldstein, S. Impression Formation in Dyads: The Temporal Dimension. In M. Davis (Ed.), *Interaction Rhythms*. Human Sciences Press Inc., 1982, 207-224.

[10] Grinter, R., and Palen, L. Instant Messaging in Teen Live. *In Proceedings of CSCW ’02*. ACM Press, NY, 2002, 21-30.

[11] Handel, M., and Herbsleb, J.D. What Is Chat Doing in the Workplace? *In Proceedings of CSCW ‘02*. ACM Press, NY, 2002, 1-10.

[12] Herbsleb, J.D., Atkins, D., Boyer, D.G., Handel, M., and Finholt, T.A., Introducing Instant Messaging and Chat into the Workplace. *In Proceedings of CHI ‘02*. ACM Press, NY, 2002, 171-178.

[13] Horvitz E., Koch P., Kadie C. M., and Jacobs A. Coordinate: Probabilistic Forecasting of Presence and Availability. *In Proceedings of UAI'02*. AAAI Press, CA, 2002, 224-233.

[14] Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D.J., and Kamm, C. The Character, Functions, and Styles of Instant Messaging in the Workplace. *In Proceedings of CSCW ’02*. ACM Press, NY, 2002, 11-20.

[15] Kraut, R.E., Fish, R.S., Root, R.W., and Chalfonte, B.L. Informal Communication in Organizations: Form, Function, and Technology. In S. Oskamp & S. Spacapan (Eds.), *People’s Reactions to Technology*. Sage, Newberry Park, 1990, 145-199.

[16] Lotus Sametime http://www.lotus.com/sametime

[17] Nardi, B., Whittaker, S., and Bradner, E. Interaction and Outeraction: Instant Messaging in Action. *In Proceedings of CSCW ’00*. ACM Press, NY, 2000, 79-88.

[18] Pennebaker, J. W., Francis, M. E., and Booth, R. J. *Linguistic inquiry and word count (LIWC)*. Erlbaum Publishers, Mahwah, NJ, 2001.

[19] Slatalla, M. The office meeting that never ends. *New York Times*, Sept. 23, 1999.

[20] Voida, A., Newstetter, W. C., and Mynatt, E. D. When conventions collide: the tensions of instant messaging attributed. *In Proceedings of CHI ‘02*. ACM Press, NY, 2002, 187-194.

[21] Whittaker, S., Frohlich, D., and Daly-Jones, O. Informal workplace communication: what is it like and how might we support it *In Proceedings of CHI ‘94*. ACM Press, NY, 1994, 131-137.

[22] Worldwide Enterprise Instant Messaging Applications 2005-2009 Forecast and 2004 *Vendor Shares: Clearing the Decks for Substantial Growth*. IDC Market Analysis, 2005.