Families on **FACEBOOK**

Reading Session 1 – Computational Social Media (DH-500)

Ravinithesh Annapureddy

Original Paper by Moira Burke, Lada A. Adamic, and Karyn Marciniak

Good afternoon all, I am Ravinithesh. I will be presenting the paper titled Families on Facebook for the first reading session of the course. This paper was originally part of 2013 AAAI proceedings. Let's start.

Families and Communication

modern
family
communication

It is a common for us to talk with our family members regularly. One of the most important family communication is between parents and children. This starts from a young age and lasts for a long time. There has been a lot of research on offline family communication. However, in recent years, social media has brought new dynamics into numerous relationships including families. Not only the young ones but parents and grandparents are using these sites. In this context, this paper examines whether social media (particularly Facebook) is changing the quality and frequency of communication with family members or not?

Parent-Child Relationships and Communication

- Communication decreases as children grow and move away (geographically).
- Mothers talk more about emotions, thoughts, and feelings with daughters than with sons.
- Both parents encourage their sons more than daughters to be independent and to control their feelings.
- Father-son relationships are the least emotionally charged ones.

EPFL

First let's see what are the conclusions about parent child communication we have based on offline data. It is said that communication frequency decrease as the distance between the parents and children increases. Then Mothers are said to be talking more about emotions, that too with daughters than Sons. Following, we have that both parents encourage the sons to be more independent and finally, the father-son relationships are least emotionally charged.

Computer-mediated Family Communication

- Facebook – for adults and kids.
- Two-thirds of parents of children aged 12-17 now use a social networking site and 80% of them have "friended" their child.
- Why are parents using social networks?
  - Monitor their teens
  - Effort to mitigate bullying and bad behavior.

EPFL

With the advent of internet and more particularly social media, it has become easy to communicate. Until Facebook happened, the family relations on social media were secretive because the kids were using fake identities and parents were also less active. However, after Facebook, parents and teens are active users. Almost 66% of parents of children aged 12-17 now use a social networking site and 80% of them are friends on Facebook. Some primary reasons why Parents use social network is to monitor their teens and to make an effort to mitigate bullying and bad behavior.

Research Questions

1. How common are parent-child relationships on Facebook?
   - How does it vary with age and gender?
2. Connections: Who "friends" whom, and when does it happen?
   - What is the composition of their mutual friends?
3. Communication: How often do parents and children communicate on the site, and how does it vary with the child's age, geographic distance, and gender?
4. How do the subject matter and linguistic properties of conversation between parents and children vary by age and gender?

EPFL

With the offline research and rich parent child relationships on Facebook. The authors try to answer the following research questions in the paper.

First, How common are parent-child relationships on Facebook? How does it vary with age and gender?
Second, Who "friends" whom, and when does it happen? What is the composition of their mutual friends?
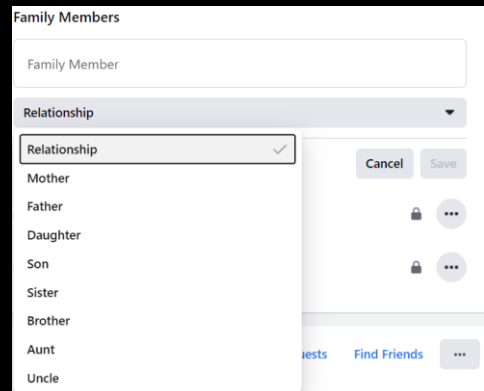Third, How often do parents and children communicate on the site, and how does it vary with the child's age, geographic distance, and gender?
And the last, How do the subject matter and linguistic properties of conversation between parents and children vary by age and gender?

Next let us look at data used for this research.

Data: Samples

- Data of US users who are
  - English speaking
  - Monthly active
  - And specified parent/child relation.
- Minimum age of parents is at least 29.

Family Members

Family Member

Relationship ▼

Relationship	✓
Mother	
Father	
Daughter	
Son	
Sister	
Brother	
Aunt	
Uncle	

Cancel   Save

🔒 •••

🔒 •••

uests   Find Friends   •••

EPFL

The users who are considered for this research were all English-speaking, monthly active US users who had specified at least one other user as their parent or child using the site's relationship tool, which is on the right side.

One restriction in creation of parent-child pairs were to remove those pairs who have less than 16 years of age difference.

Data: Features

- When was the friend request sent?
- When was it confirmed?
- When did the users join FB?
- Total number of friends, common friends.
- Communication content like posts, comments etc.
- Chats are excluded as they are biased towards people who use it.

EPFL

The features of the dataset are the time at which the friendship was initiated, when it was confirmed by the other person, when the users each joined the site, their total numbers of friends, and the friends they have in common. Also three months' communication data of content directed at any Facebook friend are also included.

Important point here is that the chats were excluded as chat volume surpasses other forms of communication but is biased toward the subset of people who use it, and the text is often too short and noisy for substantive language analysis.

Method: Pre-Processing

- Create a regression model to classify if the communication target was a family member or not.
- Text from a 1% sample of all directed communication written by the users in the study.
- Pre-processing steps
  1. Generate frequency counts for all n-grams up to three words long.
  2. Remove terms written by fewer than 1,000 people and that appeared fewer than 10,000 times were excluded.
  3. Punctuation was removed and numbers were replaced with <number>.
  4. Stop words were included as their use has been connected to many social phenomena.
     1. n-grams consisting entirely of stop words were removed.
  5. Common US first names were removed.
- 57,964 n-grams in the dictionary were generated.

EPFL

The proposed method is a language model predictive of parent-child relationships based on the content posted in Facebook. The data for this was 1% sample of all directed communication written by users. To generate the text features needed for this regression model, the following pre-processing steps are performed. First, frequency of all n-grams up to 3 words long is calculated. To avoid overfitting terms written by fewer than 1,000 people and that appeared fewer than 10,000 times were excluded. Punctuation was removed and numbers were replaced with <number>. Stop words, highly frequent words such as articles, were included as their use has been connected to many social phenomena, but n-grams consisting entirely of stop words were removed. Common US first names1 were removed. This resulted in 57, 964 n-grams at the end.

Method: Analysis

- Regression works best if features are independent
- N-grams from natural text are highly correlated.
- Elastic Net Regression
  - Combines L1 and L2 losses.
  - α controls the effect of loss.
    - $\alpha = 0$ → Ridge Regression
    - $\alpha = 1$ → Lasso Regression
- $\alpha = 0.1$ chosen for this work

EPFL

For this analysis, the model was build on logistic regression. Regression works best when features are independent, but natural language terms are highly correlated. Therefore, elastic-net logistic regression is used. The core idea of it is to combine the L1 and L2 loss which are used in the Lasso and Ridge Regression. It has a parameter, α, when it is zero it is equal to ridge regression i.e., using an L2 norm for regularization and if it is equal to 1 it is equivalent to lasso regression. After comparing the accuracy, the authors have decided to use 0.1.

Now, let us look at the results obtained.

Result: Prevalence of parent-child relationships on Facebook (RQ1)

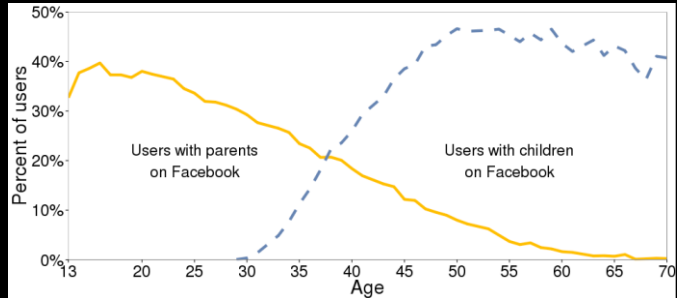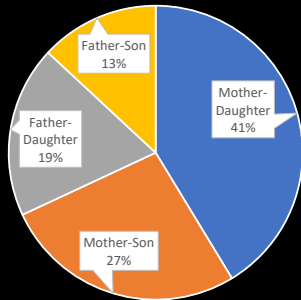The first question was to see how common are parent-child relationships on Facebook? How does they vary with age and gender?

Overall, 37.1% of English speaking, monthly-active US Facebook users have specified either a parent or child relationship on the site. Here we can see breakdown by age on the right side. Nearly 40% teens specified a parent. Almost half of the users aged 50 have a child on the site. The composition of the ties looks like this on the left. Mother-daughter ties are most common (41.4%), followed by mother-son (26.8%), father-daughter (18.9%) and father-son (13.1%).

Result: Friending (RQ2)

- Average time to friend was 371 days.
- Second person to join was more likely to send a request.
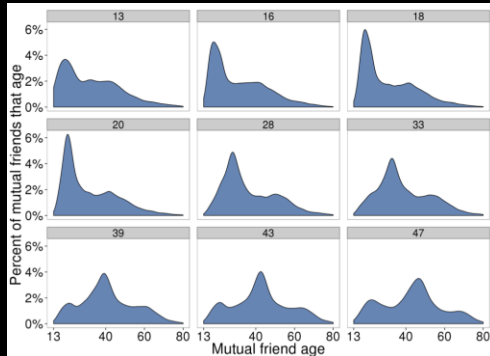- Parent – Child pairs share friends.

Figure 2 from the paper

The second question was to see who friends whom and the composition of their mutual friends?

On Facebook, the friending time varied even though the parent and children know each other offline. 19.3% of the pairs did so within a month of the second person joining, but on average, the time elapsed is 371 days. The second person to join is more likely to send the friend request, independent of the age, gender and whether the user is parent or child.

The pairs have mutual friends. The distribution of mutual friend age vs percentage of mutual friends can be seen on the right. We see that for teens and youth, most mutual friends are around the child's age, with some around the parent's age. This is consistent with the idea that parents want to be a visible on Facebook by friending some of the child's friends. Older children and their parents have mutual friends at three generations: parent, child, and grandchild.

Result: Friending (RQ2)

- Mothers and daughters have the most mutual friends.
- Percentage of mutual friends increases with age.
- Jaccard coefficient =
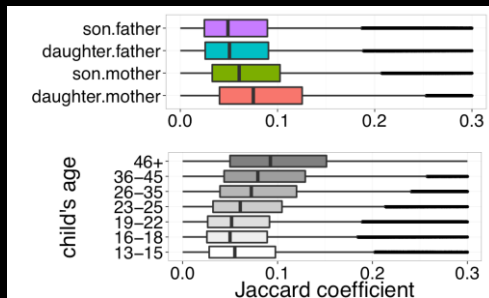$$\frac{\text{\# of mutual friends}}{\text{\# of friends}}$$

Figure 3 from the paper

EPFL

Now, comparing the mutual friends distribution among duos, From the top chart we can see that mothers and daughters have the most mutual friends, followed by mothers and sons. Although fathers have 18% more social circle than Mothers, they have fewer mutual friends.

In the bottom chart you can see that the proportion of shared contacts between parents and children also increases with age.

Result: Communication Volume (RQ3)

- Parent hold more communication to children than the opposite.
- Mother-Daughter communication rises during childbearing years.
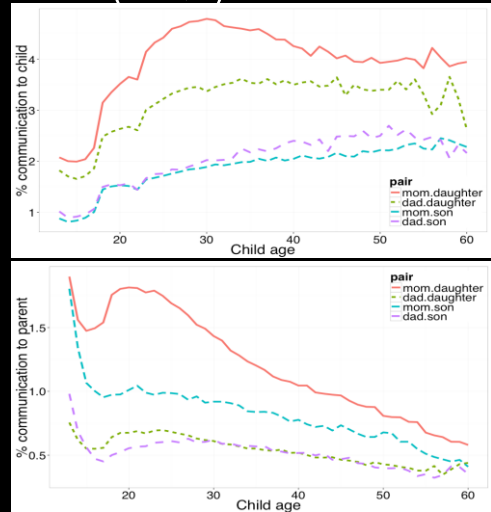- Child to Parent communication when they start moving out of the home.

Figure 4 from the paper

The third question is to see how often do parents and children communicate on the site, and how does it vary with the age, geographic distance, and gender?

From the top graph we can see that only a small percentage of directed communication from parents is to their children, with mothers writing four times as often as fathers. There is a bump between mothers and their daughters of childbearing ages, possibly where the women discuss parenting and family life.

Looking at the bottom graph, we see that children on average have 1% of their posts to their parents, and it decreases with age; There is a jump around age 18 as teens move out of the home.

Result: Communication Volume (RQ3)

- Communication frequency remain constant with distance.
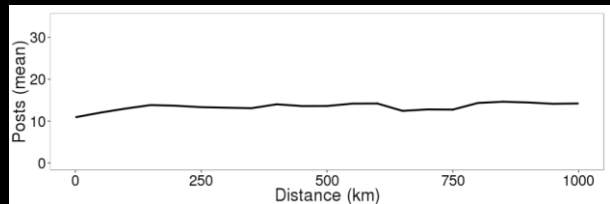- This is contrasting to prior research.

Figure 6 from the paper

EPFL

In contrast to prior studies the authors find that communication frequency remains constant with geographical distance. It is possible that the usage of Facebook itself is causing this effect.

Result: Communication Content (RQ4)

- Top terms used by parents to communicate with their children.
- Parents reference phone calls and visits and give advice to their children.
- Parents use more affection.
- Grandchildren are a common topic.

Parents writing to their children

Phrase	Odds	Phrase	Odds
love mom	3.32	gets that	2.34
week from	3.10	babygirl	2.34
that anyone	2.97	http apps	2.34
tried calling	2.89	dad	2.32
night or	2.86	http apps facebook	2.29
copy of this	2.83	carving	2.27
grammy	2.77	dr oz	2.27
home with you	2.72	has lost	2.25
papa	2.64	school work	2.23
ur sister	2.61	takes after	2.23
no thanks	2.59	taught you	2.23
it twice	2.51	love ya both	2.23
son	2.51	here we come	2.20
of them lol	2.48	your room	2.20
waking	2.41	te amo	2.18
dolly	2.39	keep working	2.18
car with	2.36	will get there	2.18
second one	2.34	your sister	2.18
you sick	2.34	favorite song	2.16
her new	2.34	gamers	2.14

Table 1 from the paper

The last question wanted to see the content of conversation between parents and children?

This table on the right shows the terms used by parents towards their children. We see that parents use more affection and endearing terms. At the same time they refer to visits, phone calls and grandchildren. The interpretation of the table is that the post containing grammy is 2.7 times as likely to be going to a child than to any other friend of the parent.

Result: Communication Content (RQ4)

- Terms divided into eleven categories as
  - Affect (e.g., love, happy, sad),
  - Family (daughter, husband, aunt),
  - Social (call, visit, share),
  - Home(bedroom, house, kitchen),
  - Ingestion (food, dish, pizza),
  - Leisure (ball, playstation, party),
  - Friend (friend, bf, bud),
  - Health (clinic, flu, pill),
  - Body (cheek, hands, spit),
  - Swearing, and
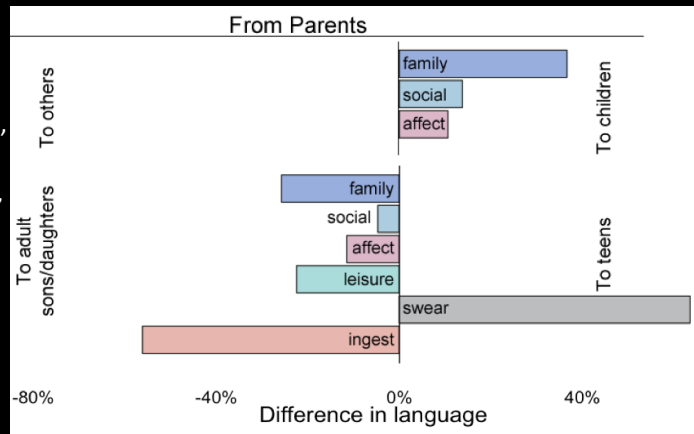 - Non-dictionary words (slang and acronyms).

Figure 7 from the paper

The authors have applied the following dictionaries on the posts, Affect, Family, Social, Home, Ingestion, Leisure, Friend, Health, Body, Swearing and other non dictionary terms. The top part of the plot on the right shows that parents use 30% more terms in the Family category when speaking to their children than to their other friends. Similarly, parents use 11.7% more Social and 9.1% more Affect words with their kids.

In case of adult children, parents use terms from the family, leisure and ingest more than the teens and use more swear terms with the teens. The latter could be an attempt from them to look cool before of their children and protect them from bullying.

Result: Communication Content (RQ4)

- Parents use words related to updates and health.

- With the teens it is food.

Parents to their teens (vs. their adult children)			
Phrase	**Odds**	**Phrase**	**Odds**
baked	4.39	on your way	2.32
mommie	3.29	who loves	2.29
looking couple	3.16	are so funny	2.27
babygirl	3.06	lo que	2.27
love mom	3.06	one way	2.25
te amo	3.00	daughter	2.25

| Parents to their adult children (age 30+, vs.			
Phrase	**Odds**	**Phrase**	**Odds**
your son	4.31	figures	2.34
titi	3.35	i totally	2.32
cousin	3.13	gota	2.29
prima	3.03	my friend	2.29
one person	2.86	division	2.27
much love to	2.72	update	2.25

Table 2 and 3 from the paper

Coming to specific terms, we see that the parents use words related to updates and health more with adult children and with the teens it is about food.

Result: Communication Content (RQ4)

- Parents treat adult children as peers .

- Parents use more affection to daughters.

- Parents push sons toward independence.

Figure 7 from the paper
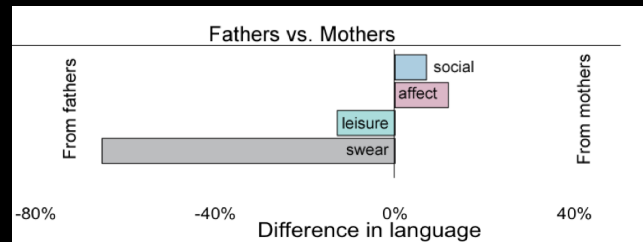
In comparing the mother and father distinction in communication, both of them appear to be treating their adult children more like peers are more likely to use affection toward their daughters rather than their sons. We see that the mothers' language is more social with their daughters, referring to other family members and fathers talk more about activities and swear a lot. As in offline communication parents appear to push their sons toward independence.

Result: Communication Content (RQ4)

- Children talk about calling and visiting.
- Virtual laughter is more in posts to parents.
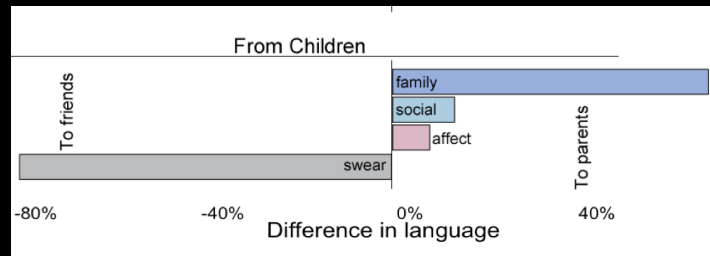
Figure 7 from the paper

Looking at the communication from children to their parents, Children talk about calling their parents and to some extent about visiting. Virtual laughter is also more predictive of posts written to parents. One interpretation is that commenting with laughter is simply a lightweight way to indicate that you've seen a post.

This finishes the answers to all four questions. Let me conclude the key points.

Conclusion

- Parent-child relationship on Facebook is similar to that exists offline.
- Children's communication with their parents decreases first and then increases.
- In contrast to previous research, the communication frequency on Facebook does not decrease with geographic distance.
- Parents initiate a lot of conversations, particularly when their daughters are raising families of their own.
- Data from English speaking US users was used.
- A regression model was built to classify if the content posted by a used is towards their parent/child or to others.
  - The features extracted from the model were used to find the defining terms of a category.
- Facebook plans to use the inferred relationships organize news feed, recommend friend connections with other family members.

EPFL

The overarching conclusion is that the parent child relation on Facebook is similar to than on the ground and decreases in the teen to youth years and then grows as the grow old. However, in contrast to one result we have seen in the introduction, the communication on Facebook remain constant with increase in geographic distance. As mentioned earlier, Facebook itself could be the reason for it. Most of the conversations are initiated by parents that to by mothers to daughters about rising families.

The Data used was from English Speaking US users and a classifier was built to detect if a post is towards parent/child or to some others. Then the coefficients of the terms used in the regression were used as defining terms for the categories. Lastly, Facebook plans to use the inferred relationships organize news feed, recommend friend connections with other family members.

Thank You

EPFL