How Social Information Networks Reflect Major Life Events: Case of Childbirth

Munmun De Choudhury   Scott Counts   Eric Horvitz

Microsoft Research, Redmond

{munmund, counts, horvitz}@microsoft.com

Social information networks, such as Twitter and Facebook provide powerful tools for individuals to share thoughts, opinions, and emotions as well as to create and maintain social ties. Considerable research in the recent past has focused on the understanding of large-scale social processes enabled by the availability of data derived from content and activity on social networks, spanning multiple domains including finance, politics, epidemiology and public health, marketing, and crisis mitigation.

Beyond exploration and inferences about population-scale dynamics and phenomena, social information networks can also provide insights at an individual scale about evolution and changes of intentions, language, mood, and behavior. As increasingly large numbers of people have been using social information networks over time scales measured in years, these networks can serve as windows onto a variety of important developments in an individual's personal life, including such events as childbirth, death of a loved one, marriage, divorce, loss of a job, and traumatic experiences such as an accident or a disaster event. From a research perspective, this allows for nuanced individual-level behavioral analyses in a longitudinal manner; over time scales long enough to include periods before and after one or more major life events.

In this light, we explore the harnessing of social information networks as a lens on behavioral changes of individuals around a particular major life event: childbirth. Having a baby is a major life event that creates significant changes in the lives of new mothers. Sleep and daily routines are disrupted, and adjustments must be made in personal and professional lives. First time mothers may be particularly challenged with navigating the new, complex realm of caring for their newborn. Adding to the challenges, many new moms experience psychological changes, such as the "baby blues," a temporary condition involving mild mood instability, anxiety, and depression. We note that whether in the form of explicit commentary, patterns of posting, or in the subtleties of language used, social media posts offer cues as to how a mother might be responding to the experience of parenting a newborn. We therefore specifically examine patterns of activity, linguistic, and emotional correlates for prenatal and postnatal periods in a set of new mothers engaged in sharing content on the social media Twitter.

At an individual scale, this method for life change measurement includes identifying and assessing emotional and behavioral expression and reactions that the individual might not even be aware of, but that could be causing very real impact on the individual's health and wellness. For instance, in the specific case of childbirth, behavioral markers, thus identified, may enable adjuvant and unobtrusive diagnosis of postnatal disorders, complementary to survey based approaches such as the Edinburgh Postnatal Depression Scale (Cox, Holden et al., 1987). They can thereby allow early intervention on the part of doctors and caregivers (e.g., psychotherapy treatments), even if a new mother is not aware of her symptoms as being out of the ordinary. At a collective scale, this research can provide improved assessment of the impact of life changes on populations. For childbirth this might mean aiding governmental agencies, support groups, or the medical community at large in the measurement of postpartum depression, traditionally an underreported disease, among large populations.

Data and Behavioral Change Measures

As we are interested in understanding behavioral change around the life event of the birth of a child, our population of interest comprises new mothers: hence we focus on female Twitter users who are likely to have given birth to a child in a particular timeframe. We follow a two-stage approach to construct this sample, described as follows.

First, we identify birth event-indicative posts on Twitter by filtering the Firehose stream based on phrases typically characterizing newspaper (The Beacon-News, Aurora News-Register, Crete News, Gothenburg Times) birth announcements: e.g., "birth, weigh*, pounds/lbs, inches, length/long, baby/son/daughter/boy/girl". The authors of these birth announcement Twitter posts thus constituted an initial set of candidate new mothers. In the second step, we identify a set of high probability new mothers, first by performing gender inference via a classifier trained on U.S. Census data, and then obtaining confidence ratings on them from crowdworkers recruited via Amazon's Mechanical Turk, in the presence of additional cues from their Twitter profiles. Our final dataset consisted of 85 validated new mothers, between May and June 2011. Finally, for each of these 85 new mothers, we queried their Twitter (public) timelines in the Firehose stream to collect all of their posts in two 5-

month periods, corresponding to prenatal and postnatal phases (Dec 1, 2010 – Apr 30, 2011 and Jul 1, 2011 – Nov 30, 2011, respectively). The bulk of our empirical analysis that follows consists of prenatal and postnatal comparisons of these 85 new mothers, including comparisons to a background cohort of 50,000 randomly selected Twitter users from the same timeframe.

We propose several measures to quantify the behavioral change of the new mothers: activity, emotion, and linguistic styles. We characterize activity via a measure we refer to as *volume*, defined as the average normalized number of posts per day made by the new mothers, over the prenatal and postnatal periods. We also focus on four measures of emotional state. The first two measures, *Positive Affect (PA)* and *Negative Affect (NA)* are computed using the psycholinguistic lexicon LIWC (http://www.liwc.net/). For the third and fourth measure, *activation* and *dominance*, we utilize the ANEW lexicon (Bradley & Lang, 1999). The approximately 2000 words in ANEW have ratings in terms of pleasure, arousal, and dominance (over a 1-10 scale); and can therefore be used to measure activation and dominance measures of Twitter posts. Finally, we introduce measures to characterize behavioral change based on the use of *linguistic styles* in posts by the new mothers. We again use LIWC for determining 22 linguistic style markers, e.g., articles, adverbs, personal pronouns, preposition, conjunction, functional words.

Empirical Studies

Prenatal Versus Postnatal Comparison. Starting with group level comparisons, Figure 1 shows a number of changes following childbirth among the 85 new mothers, compared to the background cohort. To further illustrate our observations qualitatively, we also show some randomly selected posts from mothers after childbirth in Table 1. Broadly, we find that the overall volumes of posting drops, indicating that on average women are posting less, suggesting a possible loss of social connectedness following childbirth. Within the content they do post however, we see a drop in PA and increase in NA (notice high NA in Table 1), a shift potentially attributable to the mother's physical, mental and emotional exhaustion, as well as the sleep deprivation typical of parenting a newborn.

The activation and dominance measures also drop during the postnatal phase indicating a decrease in arousal, likely because the new mothers experience overwhelming fatigue from handling daily tasks around taking care of the baby (Miller, 2002). Example posts in Table 1 further demonstrate this observation. Somewhat similarly, the use of certain linguistic styles, particularly, 1st person pronouns increases (also see Table 1), while use of 3rd person pronouns drops, possibly reflecting the emotional distancing many new mothers go through after childbirth. Though not shown here, we also observe increased use of articles, adverbs, conjunctions, swear and negation style categories during postnatal phase for our set of new mothers. Prior literature
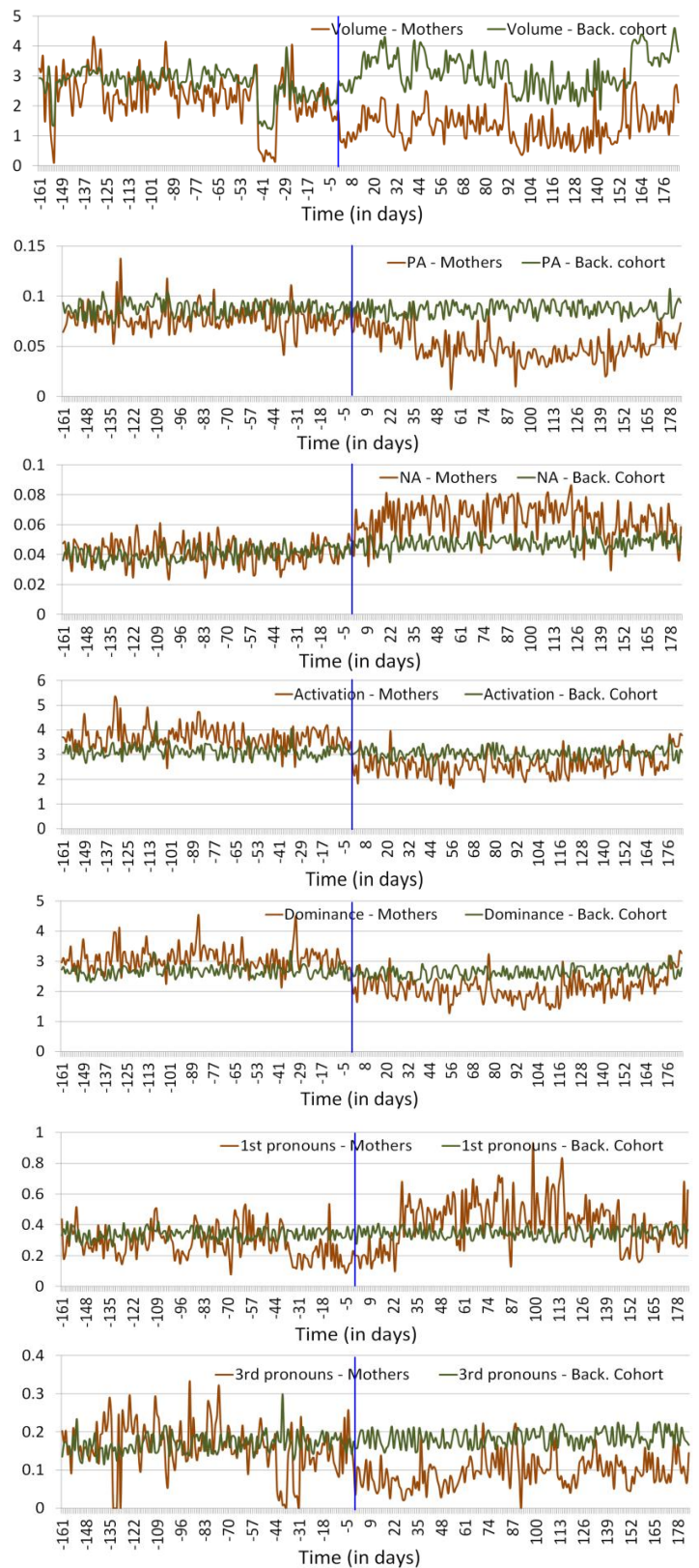

Figure 1. Comparison of behavioral measures between new mothers and the background cohort: volume, Positive Affect (PA), Negative Affect (NA), Activation, Dominance, use of 1st person pronouns, use of 3rd person pronouns.

supports high usage of these styles with expression of negative emotion, depression, or mental illness (Chung & Pennebaker, 2007) that might correspond to the circumstances of some new mothers. All of these changes are found to be statistically significant (via t-tests; $p < .001$) when comparing mothers before to mothers after childbirth (paired sample t-tests), and when comparing mothers after childbirth to the background cohort (independent sample t-tests).

Many of these changes are expected during those first few months after having a child, which is an exhausting and extremely busy time. Thus, as an initial estimate of normal change versus more extreme change, we performed a set of individual-level analyses, to determine *which* mothers in particular changed the most, and therefore what characterizes their anomalous behavior. We formalize the individual-level differences across mothers by computing Cohen's d, per mother and per measure, in order to distinguish sets of mothers with small, medium and large effect size changes following childbirth. We observe that there is a substantial number of mothers with large effect sizes for each measurement category; however activity and linguistic style measures show relatively larger number of mothers with large effect size changes. While fewer mothers undergo such changes for the emotion measures, most conservatively, we find that 12 out of 85 new mothers showed large effect size changes across *every single measure*. This set of 12 mothers then is the set of mothers whose behavior changes the most in the postnatal period across all measures, and stands out as having changed more broadly and more substantially than the other mothers studied in our data. For comparison, we perform the same exercise to determine the set of mothers who show small effects consistently across all measure types (15 mothers).

[high NA] Ugh, my daughter hates her bassinet. I hate disappointing her. What a miserable day.
[low activation] My baby is only catnapping during the day. That's so sad and depressing. I feel helpless
[low dominance] Anxiety/panic attacks need to eff off!!!!!!!!!!!!!! I'm trying to lead a somewhat normal life with my baby!!!! #frustrated #miserable
[high 1st person pronoun use] No lie I fuckin miss all socializing..... but my daughter keeps me occupied. I have all my cool moments

Table 1. Example Twitter posts from mothers with large effect size changes after childbirth.

Significant Change Postpartum. In the light of the two behaviorally distinct sets of mothers identified (large and small effect sizes), we first present a more rigorous qualitative examination of data characterizing the mothers with small and large effect sizes. We present *randomly sampled* example posts from the two cohorts in Table 2. A qualitative comparison of the nature of content shared by the two cohorts reveals that the mothers with large effects exhibit signals that are likely indicative of a lowered sense of social support ("Starting to feel lost.."), generally unhappy postings ("Feel angry n disappointed…", and even possible mental instability ("Feel like having a breakdown!", "balling my eyes out", "horrible monster mother"). Feelings expressed include anger, frustration and depression (posts (1), (3), (5)), lack of a sense of connectedness (posts (2), (3)), as well as physical discomfort and concerns about the baby (post (4)). On the other hand, the content from mothers with small effect sizes, although aligned with topics relating to bringing up the baby and expressing some sense of negativity ("Just adjusting…Need to calm down."), is less emotion-laden. For instance, we find that these mothers are using Twitter to invite comments and suggestions on their problems around typical adjustments to having a new baby–work-life balance, issues with breastfeeding etc. (posts (3), (5)).

Mothers w/ small effects
1) I know some drs say it's ok to be on meds while breastfeeding but it kind of freaks me out cause it isn't proven longterm for baby's health.
2) Days are passing by as I watch my son grow! Can't wait for more and get together with the daddy!! Wish he was here
3) Just adjusting to having a new baby, new job and we just moved town. Need to calm down. Tips/suggestions on parenting, mothers??
4) Ugh... returning to work. I'm trying to enjoy these last few days with my baby...but all I can think about is that I will be leaving him for 10 hrs a day
5) I'm taking expressed breastmilk from the fridge on outings in the diaper bag and keeping it cool with an ice pack. Someone tried it?
Mothers w/ large effects
1) This is my first baby, feel so blessed!! But angry abt being sick all the time. I guess my hormones haven't taken nicely to this big change?
2) Starting to feel lost. I'm missing my love, my baby. Feel angry n disappointed in myself. Idk what to think or do....
3) My first time being alone with my baby and I cant stop crying. What is wrong with me? Am I depressed? Im just over here balling my eyes out
4) My DS doesnt sleep more than 3 hrs at a time and cries often and is so difficult to calm down. Cant remember when was the last time I slept

5) Feel like having a breakdown! ...like the WORST mother... feel so terribly that this poor child is stuck with this horrible monster mother..

Table 2. Randomly sampled posts from mothers with small and large effect sizes.

Language Differences. Next we quantify these seemingly qualitative differences through a comparison of the overall language change (change in usage of stop word eliminated unigrams) of the set of mothers with large effects, with respect to the set of mothers with small effects, as well as the background cohort. The goal is to be able to determine *what* language accounts for the distinctive change of behavior among the mothers with large effects.

To this end, in Table 3 we show the top unigrams that exhibited the most change in usage frequency (based on Euclidian distance) between the prenatal and postnatal time periods for the background cohort, mothers that showed only small effect size changes, and finally for our set of 12 mothers who exhibited large effect size changes across all our measures.

We observe that the type of unigrams that change significantly vary substantially across the three groups. The background cohort's changes are mostly in words related to commonplace details of daily life (e.g., *tonight, here, morning, tomorrow*). For mothers with small effects, there is some evidence of going through the early childbirth phase (e.g., *fired, wait, days*). This reinforces our qualitative observation from Table 2 wherein we found these mothers using Twitter to seek support and feedback on their problems around typical baby upbringing issues. On the other hand, for the mothers with large effects, many words are emotional in nature (e.g., *aww, blessed, love*), again confirming the qualitative observations from Table 2 – see the usage of *blessed* in post (1) and the general affectionate postings (2) and (5) towards the baby.

Background cohort	Mothers w/ small effect size	Mothers w/ large effect size
now (↓), shit (↑), back (↑), that (↑), day (↓), life (↑), time (↓), them (↑), me (↑), you (↑), fuck (↑), today (↓), sleep (↑), tonight (↓), love (↓), good (↓), here(↓), ass(↓), her (↓), morning (↑), tomorrow (↑), go (↑)	#past (↑), duh (↑), people (↓), photo (↑), post (↑), decision (↓), reunite (↓), women (↑), story (↑), time (↑), asap (↓), do (↑), life (↓), wait (↑), fired (↑), days (↑), happy (↓)	haha (↓), blessed (↑), lol (↓), #lifecangetbetter (↑), awesome (↓), monthly (↑), fantastic (↓), cuddle (↑), home (↑), love (↓), sick (↑), aww (↑), scary (↑)

Table 3. Top unigrams showing the most change (in usage frequency) in the postnatal period, compared to the prenatal phase, for background cohort, mothers with small effect size, and mothers with large effect size.

Directionality of change in these words in Table 3 is critical. Considering the drops in PA and increases in NA shown earlier, along with the qualitative observations from Table 2, the observation from Table 3 that many of the changes of the emotion words are in a negative direction for the mothers with large effects, seems reasonable. For instance, use of *haha* and *lol*, frequently used terms of joviality expression in social media, are seen to drop sharply for mothers with large effect size. In fact, the example posts in Table 2 weighing heavy on negativity and social isolation make it further apparent why these mothers are not using these joviality words.

Greedy Unigram Difference Analysis. Motivated by the noticeable differences in language use among various groups, we explored the question of determining the number of unigrams whose change in usage frequencies *actually* rendered the mothers with large effects significantly different from the background cohort and those with small effects. Starting with unigrams exhibiting the most change (in usage frequency) in the postnatal phase compared to prenatal, we eliminate in a greedy iterative manner unigrams from the lexicon of all unigrams for this group, computing the Euclidian distance at each elimination step, with respect to the other two groups. As more unigrams with big change are eliminated, the Euclidian distance of language of the mothers with large effects approaches that of the other two groups. Figure 3 shows that after eliminating only 199 words (1.16% of their entire unigram vocabulary), the mothers with large effect size postnatal change are indistinguishable from those
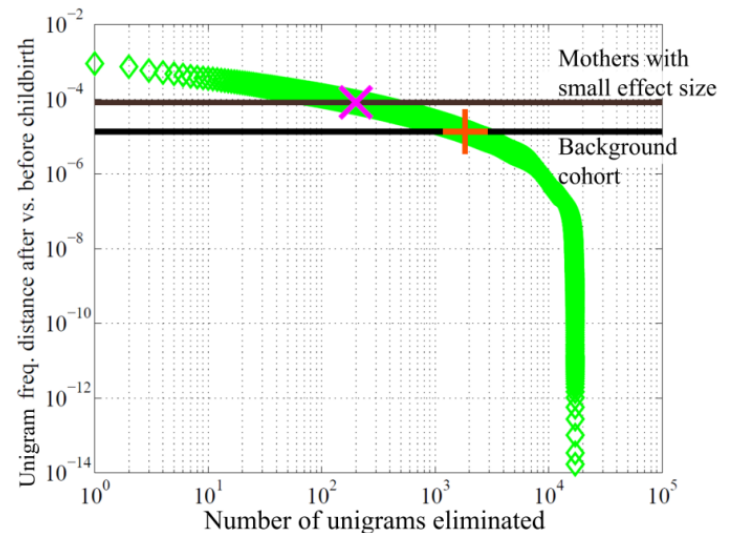

Figure 2. Unigram difference analysis technique to determine empirical thresholds that define the language change corresponding to the mothers with large effects.

with small effect changes; and eliminating 1837 words (10.73% of their vocabulary) renders them indistinguishable from the background cohort. This tells us that the changes

in the activity, emotion and stylistic measures we observed earlier appear to be subject to big changes in the usage frequencies of only a few words.

Discussion

Design Implications. Our approach and findings pave the way to several potential design implications. These include the development of automated services and tools for new mothers that can help monitor behavior and emotion in a nuanced manner, based on their social media activity. For instance, the tool could be a smartphone application that connects to the social sites the mother uses, and computes various measures over time to reveal trends in a private manner. On an individual level, monitoring some of these trends itself can serve as a self-narrative and help behavioral reflection. It could even act as an early warning mechanism to mothers with noticeably anomalous behavioral change, something especially relevant in the case of those mothers who might not be aware of their risk of PPD. By logging these trends, such an application could act as a diary-type data source to aid doctors or other trained professionals. In essence, emotional markers indicated by the tool may enable adjuvant diagnosis of postnatal disorders, complementary to survey based approaches such as the Edinburgh Postnatal Depression Scale [25], and help diagnosis or early intervention on the part of the caregivers (e.g., via psychotherapy treatments) for better health and wellness of women following childbirth.

Privacy and Ethical Considerations. Concerns regarding individual privacy, including certain ethical considerations, may arise with analyses of social media as they ultimately leverage information that may be considered sensitive, per centering on behavioral and emotional health. Our hope is that our methodology can leverage publicly available data, paired with anonymous analyses whenever possible, to generate applications that are used in complete privacy by individuals. As mentioned earlier, in our case, all data are public and, with the exception of the relatively benign Mechanical Turk task of verifying Twitter users as moms who had recently given birth, all analyses were conducted anonymously. As suggested in our design implications, the privacy of the user can be honored with user-centric design of applications that restrict the sharing of such information to the user herself and optionally to a trained medical practitioner or support group. Nevertheless, this type of research, and consequently the type of findings it generates, needs to be considered with caution, and we encourage continued discussion of the topic by the research and practitioner community.

Summary and Future Opportunities

Emotion and language are central to human functioning and social interaction, and their expressions in today's information networks potentially can be utilized to identify behavioral changes around significant life events. In this research, through a case study around childbirth, we identified a subgroup of approximately 15% of new mothers who showed broad and significant changes in behavioral and emotional expression after childbirth, a percent that aligns with published reported rates of PPD in the United States. The next step would involve validating these or similar social behavioral measures with ground truth data on PPD. For now, the narrow span of language, i.e., the 1-10% words demarcating the behavior of these 15% mothers from others, shows promise that such minimal yet discriminatory linguistic changes can be leveraged to develop language-centric diagnostic tools aiding in the identification of postpartum disorders. This could help reduce, in a non-invasive fashion, the stigma around postpartum mental illness and promote wellness among new mothers.

More generally, by way of this research, we have attempted to lay the foundation for what we believe will be a rich line of research on harnessing signals from social information networks to identify behavioral changes in individuals and for populations. Beyond core research on elucidating human behavior, such studies may lead to new kinds of methodologies for identifying when someone may need assistance, and to provide valuable and timely interventions.

References

[1] Bradley, M.M., & Lang, P.J. (1999). Affective norms for English words (ANEW). Gainesville, FL. *The NIMH Center for the Study of Emotion and Attention*.

[2] Chung, C.K., & Pennebaker, J.W. (2007). The psychological functions of function words. In K. Fielder (Ed.), *Social communication* (pp. 343-359). New York: Psychology Press.

[3] Cox J.L., Holden J.M., Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987; 150:782-6.

[4] Miller, Laura J. (2002). Postpartum Depression. *Journal of American Medical Association* (JAMA) 287 (6): 762–765.