Reality-Mining with Smartphones: Detecting and Predicting Life Events based on App Installation Behavior

Research-in-Progress

Remo Manuel Frey
ETH Zurich
Zurich, Switzerland
rfrey@ethz.ch

Runhua Xu
ETH Zurich
Zurich, Switzerland
rxu@ethz.ch

Alexander Ilic
University of St.Gallen
St.Gallen, Switzerland
alexander.ilic@unisg.ch

Abstract

Life events are often described as major forces that are going to shape tomorrow's consumer need, behavior and mood. Thus, the prediction of life events is highly relevant in marketing and sociology. In this paper, we propose a data-driven, real-time method to predict individual life events, using readily available data from smartphones. Our large-scale user study with more than 2000 users shows that our method is able to predict life events with 64.5% higher accuracy, 183.1% better precision and 88.0% higher specificity than a random model on average.

Keywords: Reality Mining, Mobile App, Life Event, Life Cycle, Life Course

Introduction

Smartphones have become our daily companions and support us in almost any situation of our daily life (Scornavacca and Barnes 2006). We constantly install, update, or delete apps on our phones to match our needs – thus making this data stream valuable for reality-mining (Chittaranjan et al. 2013; Pan et al. 2011; Ryan and Xenos 2011; Seneviratne et al. 2014; Shen et al. 2015; Trestian and Nucci 2009). It is well known that our consumption pattern differs with stages of our life (Lansing and Kish 1957; Lansing and Morgan 1955; Lydall 1955; Wells and Gubar 1966). In marketing, it is important to detect events like getting married or starting a new job in order to improve customer services or conduct more effective personalized promotion. Today, firms proactively approach their customers through surveys, emails, phone calls, advertising, or face-to-face interviews to inquire about recent or upcoming life events – which is time-consuming and not scalable. Due to the nature of the events, they are very rare and thus it is nearly impossible to detect them in a timely manner with the present methods.

In this study, we present a novel approach for detecting life events based on a smartphone user's app install behavior. We show first results from a large-scale field study with 2008 smartphone users to validate our approach. Our findings suggest that the frequency of events related to app install behavior is sufficient to perform a reliable prediction. Although the approach does not require any special permission on a user's phone, we also discuss the privacy implications and suggest an opt-in and consent based deployment of such algorithms.

The rest of the paper is structured as follows. The related work section reviews previous literature on the life events and data-driven approaches for detecting temporal changes in a user's profile. Afterwards, we introduce our research design, which is followed by a section that presents the research results. Finally, the paper concludes with a discussion of the limitations and an outlook on future work.

Related Work

Research on Life Events

In previous research, there are two predominant concepts that deal with life events. First, the concept of life cycle arose in sociology in the 1930s (Sorokin et al. 1931). Several studies (Lansing and Kish 1957; Lansing and Morgan 1955; Lydall 1955) introduced the concept into Marketing Research. A life cycle contains several life stages like having family or getting retired. During lifetime, individuals move from one stage to another, triggered by life events like having a first child or retiring. Wells and Gubar (1966) describe the importance of detecting life stages for marketing due to the fact that life stages were more meaningful than age in classifying consumers for both products and services. People in different life stages have different consumer behavior related to homeownership (Artle and Varaiya 1978), installment debt and credit card debt (Baek and Hong 2004) and different insurance needs (Blanchet 2007; Lin and Grace 2007), therefore, changes in people's life stages could lead to new business opportunities.

Another concept that is related to life events is life course, which typically refers to events, transitions, and trajectories in one's life cycle (Alwin 2012). Life courses could lead to specific consumer behavior. Rabe and Taylor (2010) did a panel study on British households and found out that couples who have a new baby are more likely to move into a better neighborhood. Similarly, numerous studies have found correlations between negative life events like divorce, job loss and death and the subsequent onset of major depression (Andrews and Tennant 1978; Costello 1982; Kessler 1997; Paykel 1978). Thus, detecting life events such as having a baby or losing a job is important because it could lead to a more effective promotion on an apartment in a good neighborhood or a timely depression treatment at an early stage.

Existing Methods for Life Event Prediction

Previous research on life event prediction typically leverages secondary data. Du and Kamakura (2006) analyzed historic panel data about the composition of members in a household in the United States for a period of 34 years and proposed a hidden Markov model to make projections about how life stages change over time, which in turn can be used to predict an individual's future life stage, income or probability of having positive expenditure on durables. Guzzo (2006) analyzed the relationship between life events and union formation and found that experience of events near union formation might provide clues to why the union formed and have implications for future outcome and stability. Yang (2009) observed and analyzed data from panel interviews and surveys about consumption of housing and non-housing goods and developed a quantitative, dynamic, and general equilibrium model to predict life-cycle behavior as well as to generate consumption profiles.

Although promising, the predictors used in those studies are consumption data, household and other demographic data, which are difficult to detect until being measured in questionnaires or interviews. However, due to the fact that life events are transitions and do not last long, detecting such events before they are taking place becomes extremely important. Research shows that life events are predictable timely based on one's purchase behavior through analyzing her credit card records (Del Bene et al. 2010). In addition, by utilizing location telematics data, it is also possible to predict life events in real-time. Examples include, but are not limited to: If a person travels to properties that are for sale with certain degree of frequency, analysis engine may determine that she is in the market for new home and may determine that targeted advertising or products (e.g. home inspection, mortgage products) related to a new home sale may be of interest to her (Healy et al. 2014).

Nonetheless, one limitation of the above-mentioned real-time prediction approaches is that the data used in those studies is only available to credit card providers, vehicle manufacturers or telecommunication companies. As data is becoming important and could also be part of a firm's core competence, it is difficult for other entities like solution providers to access such data and provide service upon that, which might make those solutions limited useful in reality.

Prediction based on Accessible Data

The proliferation of social media makes it possible for real-time life events prediction based on communications between the target user and other users of a social networking system (Paul 2013). Paul proposed an idea of sending gift suggestions to one or more friends if the target user is expecting a life event. De Choudhury et al. (2013) studied life events from 376 mothers as an independent variable and combined it with their activities on Twitter. Engagement, Emotion, Linguistic Style and other data based on Twitter posts were analyzed to forecast future changes in emotion and behavior of new mothers.

As Smartphones are owned by individuals (Scornavacca and Barnes 2006) and the use of mobile apps could be correlated with individuals' needs, interests, habits, and personality (Ryan and Xenos 2011; Seneviratne et al. 2014; Shen et al. 2015), researchers recently leveraged acoustic measurements (Pianesi et al. 2008), email content (Shen et al. 2013), face recognitions (Han et al. 2014), phone logs (Chittaranjan et al. 2013; Trestian and Nucci 2009), and social network content (Chin and Wright 2014; Minamikawa et al. 2012) to automatically predict a person's age, gender, and personality. In addition, Seneviratne et al. (2014) showed the possibility to use a snapshot of apps a user installed to predict more personal traits like religion, relationship status, and spoken languages. Pan et al. (2011) developed sensing software to collect Smartphone meta-data like phone call logs, number of nearby devices, and apps installed. Based on a sample of 200 students, the authors were able to leverage Smartphone meta-data to accurately predict events like what mobile apps a user would install in the near further.

Nowadays the number of available mobile apps in major app stores easily exceeds one million – providing an app for almost any situation of our life (Statista 2014). Based on what apps a user had installed, previous research showed the possibility to accurately predict some parts of her traits and behavior. Therefore, we suspect that it is also possible to detect and predict an individual's life events at real-time by leveraging easy accessible data such as what mobile apps she has installed to overcome the feasibility and scalability problems of current survey-based approaches.

Research Design

Research Question and Methodology

Research on detecting and predicting life events in a data-driven and non-intrusive way is scarce. Previous research is based on non-scalable methods like questionnaires, interviews and historic data. Due to the lack of real-time data, the research focus was on snapshots of life stages instead of transition probabilities from one stage to the other. New technologies in telecommunication and mobile devices enable us to predict life events with real-time data. However, to the best of our knowledge, no research has leveraged such data to conduct prediction. Thus, our central hypothesis is that app installation behavior produces enough data points for a better than random prediction of life events. Our key data source is called app installation logs that we define as a user's history of time stamped events for each app installation. Our goal is to explore the potential of this data in life events prediction. Therefore, we aim to address the following research question:

***RQ:** Is it possible to detect and predict life events based on the app installation behavior?*

To answer the research question, we use the following methodology. For each participant, we collect her app installation logs as well as a questionnaire to elicit the occurrence of a life event in the last six months and upcoming six months. We use the questionnaire for training our prediction model and for validating it as shown in Figure 1.

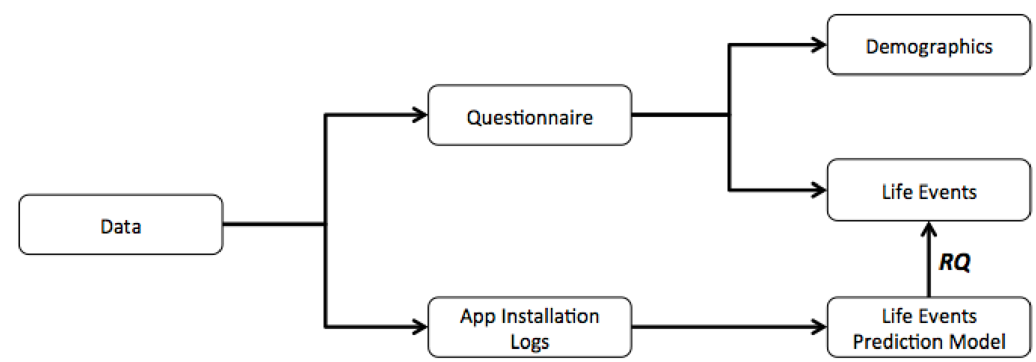


Figure 1. Overview of Research Design

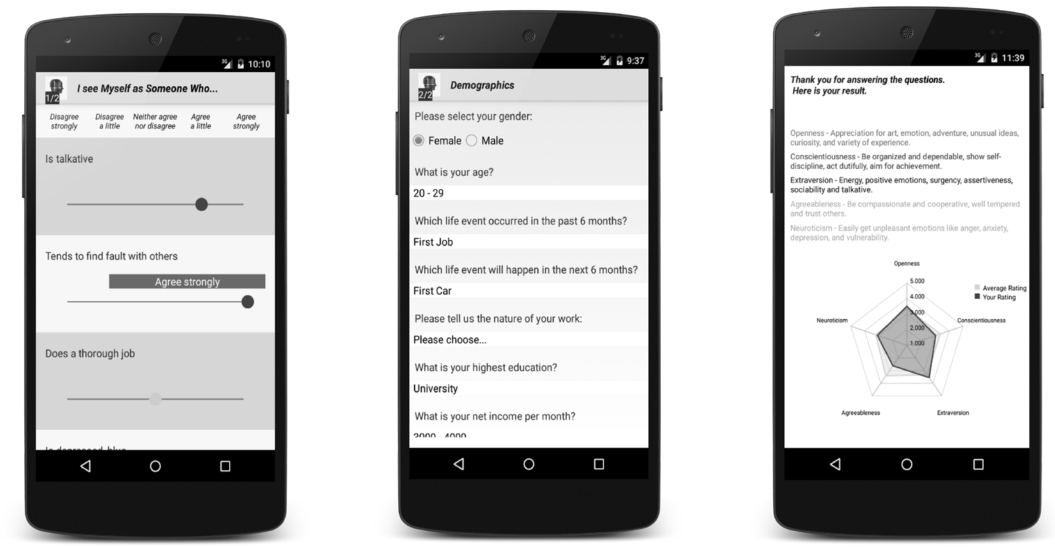
Prediction Model

We divide the sample randomly into two sets of 50% for training and 50% for test to measure our model's prediction power. In the current work, we keep our prediction model simple: To train the classifiers, we search manually in the training set for apps related to a specific life event based on app names and installed by our positive cases (i.e. participants *with* a life event). Then, we extract manually up to four of the most frequent keywords from the names of these apps. There is no distinction between upper and lower case. If one of the keyword is contained in an app name of participant's apps, then she is predicted as belonging to the positive class, else to the negative one.

Data Collection Method

Data is collected together in context of another large-scale field study, where an Android app was developed. Figure 2 shows the three main pages of the app. The app is described as a personality test game where users give answers to the personality measurement (as shown in Figure 2.a) and demographics (as shown in Figure 2.b) to compare her personality traits with the average of other people who have already participated in the game (as shown in Figure 2.c). Our two questions about life events are integrated in the demographic questions.

When the app is opened for the first time, a random and unique string is generated to represent each participant. Meanwhile, a background process in the app is initiated, which reads app installation logs from the device and sends it to a backend webserver. Once all the questions that measure personality or demographics are answered, the answers are transmitted immediately to our server. In addition, after going to the next page, it is not possible to go back to change answers. It is also impossible to redo the personality test on the same device more than once. By these restrictions, we try to prevent users from providing their own device to others who also want to do the test. The app is listed on Google Play Store and we leverage Facebook pages, news feeds and posts to encourage users to download the app.



a) Big Five Question Battery

b) Questions about Life Events and Demographics

c) Results in Star Chart

Figure 2. Screenshots of the Survey App

Data Quality Checks

Due to the fact that many smartphone manufacturers pre-install mobile apps before selling phones, such pre-installed apps are not suitable as predictors and thus should be removed. In addition, we remove any events in the app install log older than one year to ensure our algorithm is working on recent data. For this pre-study, we are primarily interested on the general occurrence of the life event in a timeframe of around 6 months of the data collection. Future work will include a higher temporal resolution and also adaption. Participants who did not answer all life events related questions were not be considered for model building and validation.

Results

Collected Samples

The personality test app was published on Google Play Store in three German-speaking countries (Germany, Austria, and Switzerland) on March 27, 2015. The corresponding Facebook feeds and posts were distributed between March 27, 2015 and April 1, 2015. During this period, our app promotion page was shown to 107,504 people and 2092 people installed the app. The conversion rate for installation is around 2%. Among the 2092 people who installed the app, 2008 people filled out the questionnaire and 1792 people answered both of the life event questions – i.e. related to past and to future life events. The survey completion rate is around 96%. The total number of distinct apps is 11,420 or 10,884 after removing the preinstalled apps. The data on demographics was not used in the model and are just provided for illustration purposes of our sample population only as seen in Table 1.

Life events are rare by definition and the results of our questionnaire in Table 2 reveal the situation with one column showing the percentage of life events happened in the past half a year and another column showing that for the next half a year. More than a half of the participants (52.5%) did not report any life event. The most frequent life event is having the first job (14.0% and 16.3%), followed by purchasing the first apartment (12.0% and 9.1%) and the first car (5.0% and 5.4%). The least frequent life event is getting retired: Less than 1% of the people in our sample are retired or will experience it in the next six months.

Table 1. Characteristics of Participants in the Study (N=2008)

Respondents	*Range*	*In %*		*Respondents*	*Range*	*In %*
Gender	Female	77.1%		Net Monthly Salary (€)	> 5000	0.4%
	Male	22.9%			4000-4999	0.4%
Age	10-19	21.7%			3000-3999	1.6%
	20-29	50.9%			2000-2999	5.2%
	30-39	17.9%			1500-1999	13.2%
	40-49	7.8%			1000-1499	23.4%
	50-59	1.4%			500-999	23.2%
	No Answer	0.3%			< 500	17.9%
Job Type	Permanent Job	42.7%			No Answer	14.7%
	Temporary Job	0.8%		Highest Education	University	5.2%
	Self-employed	2.4%			High School	14.8%
	Student	22.4%			Vocational Education	29.1%
	Seeking for Jobs	6.6%			Secondary School	44.9%
	Housewife	10.0%			Elementary School	1.4%
	Retired	0.7%			No Degree	4.6%
	No Job	1.5%				
	No Answer	12.9%				

Table 2. Results of the Life Event Questions (N=2008)

	First Car	First Job	Marriage	First Apartment	First Child	House Purchase	Divorce	Retirement	Heritage	None of the Listed Events	No Answer
Which life event occurred in the past six months?	5.0%	14.0%	1.8%	12.0%	3.7%	1.8%	2.1%	0.3%	0.6%	52.5%	6.2%
Which life event will happen in the next six months?	5.4%	16.3%	4.2%	9.1%	2.1%	1.8%	1.3%	0.1%	-	52.5%	7.4%

Resulting Model Parameter

We trained our classification model for five life events: first car, first job, marriage, first apartment and first child. The other four life events had to be dismissed since the N of the positive cases was too small to derive a reasonable training set. Since the study was conducted in German-speaking countries, most of the trained keywords are in German. In Table 3, we display the English translation of the keywords.

Table 3. Lists of Keywords for each Life Event (N=896)

First Car	First Job	Marriage	First Apartment	First Child
driver's license, driving school, learn to drive	job, job market, apprenticeship	wedding, marriage, love	real estate	pregnant, baby

Performance Evaluation of the Model

The performance of our model is shown in Table 4. Accuracy, precision, recall, specificity and positive likelihood ratio (LR+) are displayed. For comparison, we used a random model that classifies each participant positively with a probability of 50%. Another quite common model classifies all participants to the dominant class. The model improves the accuracy, but it is useless for predicting life events because no life event would be ever predicted by it.

Table 4. Life Events Classifier Performance (N=896)

(Improvement compared to random classification is displayed in brackets.)

	Accuracy		Precision		Recall		Specificity		LR+
First Car	0.867	(+73.4%)	0.229	(+104.8%)	0.080	(-84.0%)	0.966	(+93.2%)	2.4
First Job	0.685	(+37.1%)	0.358	(+23.9%)	0.112	(-77.6%)	0.918	(+83.7%)	1.4
Marriage	0.878	(+75.7%)	0.097	(+47.0%)	0.102	(-79.7%)	0.933	(+86.6%)	1.5
First Apartment	0.746	(+49.1%)	0.372	(+67.4%)	0.211	(-57.8%)	0.898	(+79.6%)	2.1
First Child	0.935	(+87.1%)	0.500	(+672.4%)	0.224	(-55.2%)	0.984	(+96.9%)	14.4

Typically, there is a tradeoff between precision, recall and specificity (Hastie et al. 2009; James et al. 2014). Our model reaches a higher accuracy (up to 87.1%), precision (up to 672.4%) and specificity (up to 96.9%) than the random model, but the recall is up to 84.0% lower. In the context of life events, a high precision is important especially for banks and insurances. Advising customers in an expensive one-to-one conversation is still common. Thus, targeting potential customers with high precision becomes crucial. High precision helps to identify as many potential sales as possible while minimizing marketing cost.

The life event 'First Child' reaches the highest values for accuracy (93.5%), precision (50.0%), recall (22.4%) and specificity (98.4%). It is understandable because of the fact that there are many useful apps especially for that life event, like apps that support pregnancy, track child growth, etc. The good result is also reflected in the high positive likelihood ratio. A ratio greater than 1 means that people who are positive classified, have an increase probability that the life event happens. The life event 'First Child' with a ratio larger than 10 has even a "large and often conclusive increase in the likelihood", according to Riddle and Stratford (1999).

To sum up, it is possible to target people with a higher accuracy, higher precision and higher specificity compared to our random model. Furthermore, statements about prediction of individuals' life events are better than random in general. Thus, the research question is answered.

Discussion, Limitations, and Future Work

By examining app installation logs of 2008 Android phone users, we generate a model to predict each smartphone user's life events. On average, our prediction model perform 64.5%, 183.1%, and 88.0% better than random guess in terms of model accuracy, precision and specificity, respectively. However, the recall rate is on average 70.8% lower. Our model predicts extremely well for the life events like 'First Car'.

The contribution of our work is two-fold. First, life events prediction enables firms like insurance companies and banks to identify potential customers in real-time. Due to the fact that people do not change insurance and financial companies frequently (Statista 2012), it becomes crucial for such companies to address customers before they choose a product from a competitor. Take insurances for example, it is regularized in most western countries that insurance is required for a new car, a new house or apartment, and an infant. If life events like buying a new car, a new apartment, and giving birth are predictable, firms are able to conduct more effective personalized promotion (Dorotic et al. 2012) and cross-selling to the right people at the right time thereby increasing the customer conversion rate and locking in the customers over a long time. Second, state-of-the-art approaches of detecting or predicting life events are typically based on paper questionnaires or personal interviews, which is time-consuming and limited scalable (Montjoye et al. 2013). By integrating our approach in their existing mobile apps, firms can predict customers' next life events without bringing the pains of answering questionnaires or conducting interviews. Compared with other data-driven prediction approaches proposed by previous research, our approach has no limitation on data access. Technically, all solution providers or individual developers can retrieve the data through Google API.

Although powerful, both retrieving mobile app data and conducting personalized marketing might trigger users' concern about privacy (Chen and Hsieh 2012; Lam et al. 2006). We suggest app publishers that leverage the approach to state explicitly to corresponding app users regarding information like when and what data will be collected and for what purpose. Each well-designed app should be transparent on data collection. App publishers should also give users the right to opt-in for providing the mobile app installation data and receiving personalized in-app recommendations and promotions, according to the suggestion from Pentland (2014). However, compared to existing data-driven approaches that trace the installation of specific apps and phone call logs for user profiling, our prediction model should lessen privacy concerns because only the app name instead of app usage time is used in our approach.

There are several limitations of this paper. First, although we have a large sample size compared with previous research, the sample is not representative in terms of age, gender, and salary. Specifically, more than 75% of our sample are women and more than 70% of the sample are younger than 30 years old. This is understandable as we distributed the app on Facebook and previous research shows that women are more likely to heavily use Facebook and the like (Raacke and Bonds-Raacke 2008; Thompson and Lougheed 2012). Future research will aim to specifically address a more balanced sample selection to explicitly study the impact of demographics on the performance.

Second, the current work uses a very simple model that is based on app name key words to predict life events. Although the precision is already quite good, the resulted recall rate could be further improved. We suspect strongly that the key words used in our model are not representative for most of the participants who have certain life events. Therefore, more sophisticated machine learning methods like Support Vector Machines or Neuronal Network should be used in future research to automatically find out other key words that are more predictive but less intuitive. Besides, as app names are given by the developers, it could happen that developers will follow new trends and use other app names in the future. Therefore, whether and how frequently our prediction model needs to be retained requires future studies.

Third, life events are typically time-dependent. In this work, we predicted a user's life events in a one-year time-window (six months before and after the survey) based on the apps she has installed in a one-year time-window (in the past year). However, both time-windows are not ideal and by varying the length of them, prediction models and/or prediction accuracy could change. Thus, different length of both time-windows needs to be tested to get a better understanding and reasonable results.

Furthermore, due to the limited space we have in the survey app, we only used a drop-down list for participants to select one out of the eight defined life events. But it is possible that a participant might have experienced or will experience more than one life event at the same time. In addition, due to the lack of response data, we did not generate prediction models for the other four life events. Future research should collect more data about these events and test the possibility for predicting them.

Conclusion

We presented a novel method to detect and predict life events by analyzing the app installation behavior of smartphone users. The results from our large-scale study clearly indicates that our data-driven approach is a useful instrument for practitioners and researchers as well. For practitioners, the automated and scalable nature of our approach could open up opportunities for more pro-active service provision and novel business models. Researchers in life cycle and life course research are able to study life events in real-time and receive a lightweight alternative to intrusive questionnaires and interviews.

Acknowledgments

We thank our research partner 42matters AG in Switzerland who supported us with knowledge and experience in the field of mobile technologies.

References

Alwin, D. F. 2012. "Integrating varieties of life course concepts," *Journals of Gerontology - Series B Psychological Sciences and Social Sciences* (67 B:2), pp. 206–220.

Andrews, G., and Tennant, C. 1978. "Life event stress and psychiatric illness," *Psychological medicine* (8:4), pp. 545–549.

Artle, R., and Varaiya, P. 1978. "Life cycle consumption and homeownership," *Journal of Economic Theory* (18), pp. 38–58.

Baek, E., and Hong, G. S. 2004. "Effects of family life-cycle stages on consumer debts," *Journal of Family and Economic Issues* (25:3), pp. 359–385.

Del Bene, M., Hu, P., Nambiar, A., and Salazar, D. G. 2010. "Methods and systems for predicting consumer behavior from transaction card purchases," U.S. Patent 20100161379.

Blanchet, D. 2007. "Changing life cycles and their impact on insurance," *The Geneva Papers on Risk and Insurance Issues and Practice* (32), pp. 1–10.

Chen, P. T., and Hsieh, H. P. 2012. "Personalized mobile advertising: Its key attributes, trends, and social impact," *Technological Forecasting and Social Change* (79:3), Elsevier Inc., pp. 543–557.

Chin, D. N., and Wright, W. R. 2014. "Social media sources for personality profiling," in *Proceedings of the 2nd Workshop Emotions and Personality in Personalized Services (EMPIRE 2014)*, Aalborg, Denmark.

Chittaranjan, G., Blom, J., and Gatica-Perez, D. 2013. "Mining large-scale smartphone data for personality studies," *Personal and Ubiquitous Computing* (17:3), pp. 433–450.

De Choudhury, M., Counts, S., and Horvitz, E. 2013. "Predicting postpartum changes in emotion and behavior via social media," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, Paris, p. 3267.

Costello, C. G. 1982. "Social factors associated with depression: a retrospective community study," *Psychological Medicine* (12:02), p. 329.

Dorotic, M., Bijmolt, T. H. A., and Verhoef, P. C. 2012. "Loyalty programmes: Current knowledge and research directions," *International Journal of Management Reviews* (14:3), pp. 217–237.

Du, R. Y., and Kamakura, W. a. 2006. "Household life cycles and lifestyles in the United States," *Journal of Marketing Research* (43:February), pp. 121–132.

Guzzo, K. B. 2006. "The relationship between life course events and union formation," *Social Science Research* (35), pp. 384–408.

Han, H., Otto, C., Liu, X., and Jain, A. 2014. "Demographic estimation from face images: Human vs. machine performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (37:6), pp. 1148 – 1161.

Hastie, T., Tibshirani, R., and Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second Edi.), Springer New York.

Healy, J. J., G. Canario, P., Drawert, S. T., M. Brown, C., Gonzales, R., Dohner, R. K., and Camarano, J. 2014. "Systems and methods for utilizing telematics data," U.S. Patent 20140214750.

James, G., Witten, D., Hastie, T., and Tibshirani, R. 2014. *An Introduction to Statistical Learning: With Applications in R* (G. Casella, S. Fienberg, and I. Olkin, eds.), Springer.

Kessler, R. C. 1997. "The effects of stressful life events on depression," *Annual review of psychology* (48), pp. 191–214.

Lam, S. K. T., Frankowski, D., and Riedl, J. 2006. "Do you trust your recommendations? An exploration of security and privacy issues in recommender systems," in *International Conference on Emerging Trends in Information and Communication Security*, Freiburg, Germany.

Lansing, J. B., and Kish, L. 1957. "Family life cycle as an independent variable," *American Sociological Review* (22:5), pp. 512–519.

Lansing, J. B., and Morgan, J. N. 1955. *Consumer Finance over the Life Cycle* (L. H. Clark, ed.) (Vol. 2.), New York: New York University Press.

Lin, Y., and Grace, M. F. 2007. "Household life cycle protection: Life insurance holdings, financial vulnerability, and portfolio implications," *Journal of Risk and Insurance* (74:1), pp. 141–173.

Lydall, H. 1955. "The life cycle in income, saving, and asset ownership," *Econometrica: Journal of the Econometric Society*, pp. 131–150.

Minamikawa, A., Fujita, H., Hakura, J., and Kurematsu, M. 2012. “Personality estimation application for social media,” *The Frontiers in Artificial Intelligence and Applications*, IOS Press.
Montjoye, Y. De, Quoidbach, J., and Robic, F. 2013. “Predicting personality using novel mobile phone-based metrics,” in *Social Computing, Behavioral-Cultural Modeling and Prediction*, Springer Berlin Heidelberg, pp. 48–55.
Pan, W., Aharony, N., and Pentland, A. 2011. “Composite social network for predicting mobile apps installation,” in *Proceedings of The Twenty-Fifth AAAI Conference on Artificial Intelligence*, San Francisco, California, pp. 821–827.
Paul, A. 2013. “Targeting items to a user of a social networking system based on a predicted event for the user,” U.S. Patent 20130159115.
Paykel, E. S. 1978. “Contribution of life events to causation of psychiatric illness,” *Psychological medicine* (8:2), pp. 245–253.
Pentland, A. 2014. *Social Physics: How Good Ideas Spread - the Lessons from a New Science*, New York: Penguin Press.
Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., and Zancanaro, M. 2008. “Multimodal recognition of personality traits in social interactions,” *Proceedings of the 10th international conference on Multimodal interfaces - IMCI ’08*, New York, New York, USA: ACM Press, p. 53.
Raacke, J., and Bonds-Raacke, J. 2008. “MySpace and Facebook: Applying the uses and gratifications theory to exploring friend-networking sites,” *Cyberpsychology & Behavior* (11:2), pp. 169–174.
Rabe, B., and Taylor, M. 2010. “Residential mobility, neighbourhood quality and life-course events,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* (173:3), pp. 531–555.
Riddle, D. L., and Stratford, P. W. 1999. “Interpreting validity indexes for diagnostic tests: an illustration using the Berg balance test,” *Physical therapy* (79), pp. 939–948.
Ryan, T., and Xenos, S. 2011. “Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage,” *Computers in Human Behavior* (27:5), Elsevier Ltd, pp. 1658–1664.
Scornavacca, E., and Barnes, S. J. 2006. “Barcode enabled m-commerce: strategic implications and business models,” *International Journal Mobile Communications* (4:2), pp. 163–177.
Seneviratne, S., Seneviratne, A., Mohapatra, P., and Mahanti, A. 2014. “Predicting user traits from a snapshot of apps installed on a smartphone,” *ACM SIGMOBILE Mobile Computing and Communications Review* (18:4), pp. 1–8.
Shen, J., Brdiczka, O., and Liu, J. 2013. “Understanding email writers: Personality prediction from email messages,” in *The 21st International Conference on User Modeling, Adaptation and Personalization*, Rome, Italy, pp. 318–330.
Shen, J., Brdiczka, O., and Liu, J. 2015. “A study of Facebook behavior: What does it tell about your Neuroticism and Extraversion?,” *Computers in Human Behavior* (46:1), Elsevier Ltd, pp. 32–38.
Sorokin, P., C., Z. C., and J., G. C. 1931. *Systematic Source Book in Rural Sociology*, Minneapolis: University of Minnesota Press.
Statista. 2012. “Willingness of changing the main bank,” in *Statista*, Visited on May 15, 2015, p. http://www.statista.com/statistics/248044.
Statista. 2014. “Number of apps available in leading app stores as of July 2014,” in *Statista*, Visited on March 15, 2014, p. http://www.statista.com/statistics/276623.
Thompson, S. H., and Lougheed, E. 2012. “Frazzled by Facebook? An exploratory study of gender differences in social network communication among undergraduate men and women,” *College Student Journal* (46:1), pp. 88–99.
Trestian, I., Nucci, A., Ranjan, S., and Kuzmanovic, A. 2009. “Measuring serendipity: connecting people, locations and interests in a mobile 3G network,” in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference - IMC ’09*.
Wells, W., and Gubar, G. 1966. “Life cycle concept in marketing research,” *Journal of Marketing Research* (Ill:November), pp. 355 – 363.
Yang, F. 2009. “Consumption over the life cycle: How different is housing?,” *Review of Economic Dynamics* (12:3), Elsevier Inc., pp. 423–443.