A State Transition Model for Mobile Notifications via Survival Analysis

Abstract

Mobile notifications have become a major communication channel for social networking services to keep members informed and engaged. As more mobile applications push notifications to members, they constantly face decisions on what to send, when and how. A lack of research and methodology commonly leads to heuristic decision making. Many notifications arrive at an inappropriate moment or introduce too many interruptions, failing to provide value at the right time and spurring members' complaints. In this paper we explore unique features of interactions between mobile notifications and member engagement. We propose a state transition framework to quantitatively evaluate the effectiveness of notifications. Within this framework, we develop a survival model for badging notifications assuming a log-linear structure and a Weibull distribution. Our results show that this model achieves more flexibility for applications and superior prediction accuracy than a logistic regression model. In particular, we provide an online use case on notification delivery time optimization to show how we make better decisions, drive more member engagement, and provide more value to members. The paper has been submitted to WSDM 2019.

1 Introduction

Social networking services (e.g., LinkedIn, Facebook, Instagram, Twitter, WeChat) actively push information to their members through mobile notifications. As the content ecosystem and members' connection network grow, more and more information is generated on the social networking site that is worth informing the members. On the other hand, members have limited attention span, regardless of how much value notifications could inform them of. The discrepancy between increasing content and limited member attention is the challenge many mobile applications are facing, especially those social networking applications.

Compared with email communication, mobile notifications are more time sensitive and responded more promptly [2, 7]. Without an established way to determine delivery time, mobile notifications often arrive at inconvenient moments, failing to draw a member's attention. Moreover, due to the pervasive nature of smartphones, such inconvenience may lead to complaints or even disablement on future notification deliveries, causing a permanent loss to both service providers and members. In short, sending notifications at the right time with the right content in many cases is critical.

The interaction of a member with mobile notifications can be very complex and depends on numerous aspects [5, 6, 9]. It is common to link a notification event to one or more rewards to evaluate the effectiveness of a notification. For engagement, a typical reward is a visit from the member to the app. One challenge for such a study is how to attribute a reward, because members may receive multiple notifications before they open and visit the app. Simple strategies could be to attribute the reward to the most recent one or to several notifications delivered within a look-back time period. Such strategies are hard to justify theoretically and could introduce significant bias in learning. Our strategy is to leverage the survival analysis to attribute a reward without ambiguity and bias [1, 3].

In this paper, we develop a state transition model to quantify how effective mobile notifications are at bringing value to the member through a visit as a reward. We focus on badging notifications, which has a more subtle effect to model and is less studied in the literature. In addition, badging notifications usually account for the majority of the total send volume. The methodology can be extended to UI push notifications with possibly different distribution assumptions as they are responded to more quickly.

Let M be a notification event, s be a mobile context state, and t^s be the time to the next visit since the start time of the state s. After a notification M_0, a mobile context state stays at s_0. Then at any evaluation time, we consider whether or not to send a notification M_1 to a member, who has stayed in state s_0 for W_0 time. The mobile context state will change to s_1 if M_1 is received. Note that a member's visit can also change the state, so s_0 may start from the most recent visit event or the most recent notification event, whichever comes later.

In our state transition model, we assume that members' engagement behaviors depend on both their mobile context states and members' characteristics. If M_1 is sent, then state s_1 kicks in and the probability of a member visiting our app within the next T time would be

$$\Pr(t^{s_1} < T \mid \boldsymbol{z}, s_1) = F_{t^{s_1}|\boldsymbol{z},s_1}(T), \tag{1}$$

where s_1 represents this member's mobile context state and $\boldsymbol{z}$ represent this member's features and $F_{t_{s_1}|\boldsymbol{z},s_1}$ is the cumulative distribution function of time-to-visit t_{s_1} given $(\boldsymbol{z}, s_1)$. T is the prediction window whose value is usually chosen based on the specific problem instance.

If we decide not to send a notification M_1, the member will stay in the current state s_0. Then the probability of the next visit within the next T time is

$$\begin{aligned} & \Pr(t^{s_0} \le T + W_0 | \boldsymbol{z}, s_0, t^{s_0} > W_0) \\ = & \frac{\Pr(W_0 < t^{s_0} \le T + W_0 | \boldsymbol{z}, s_0)}{\Pr(t^{s_0} > W_0 | \boldsymbol{z}, s_0)} \\ = & \frac{F_{t^{s_0}|\boldsymbol{z},s_0}(T + W_0) - F_{t^{s_0}|\boldsymbol{z},s_0}(W_0)}{1 - F_{t^{s_0}|\boldsymbol{z},s_0}(W_0)}. \end{aligned} \tag{2}$$

We name the difference between (1) and (2) the Δ effect, which is a function of T and W_0 given $\boldsymbol{z}$, s_0 and s_1. The larger the Δ effect is, the more value we can bring to a member by sending a notification.

One of the challenges for learning $F_{t|\boldsymbol{z},s}(T)$ is that we can not always observe the time to visit after a notification send event, because we may send out another notification before the member's next visit. With a common existence of such censoring in observational mobile notification data, we train an accelerated failure-time (AFT) model [8, 4] to estimate the the distribution of members' time to visit $F_{t|\boldsymbol{z},s}(T)$. We assume a Weibull distribution, which leads to a monotonic Δ effect over time T. Our offline evaluation shows that this survival model achieves favorable prediction accuracy compared with non-survival regression approaches.

We provide two use case formulations for notification volume optimization (VO) and delivery time optimization (DTO) separately. We then present an online use case on notification DTO, where our model is used to make send decisions. The A/B test results show significant improvement on member engagement and content consumption over a non-DTO control and a baseline DTO model. Table 1 shows the A/B test results comparing the DTO based on our model with the control and the baseline models. The experiment was tested over a full week and the numbers in the table are all statistically significant. These positive numbers validate the effectiveness of our model for DTO.

Table 1: Online A/B test results for delivery time optimization

Metric	vs. Control	vs. Baseline
Sessions	+ 1.86%	+0.67%
Engaged Feed Sessions	+ 1.78%	+0.69%
Notification Sessions	+6.19%	+1.51%
Notification Daily Unique Send CTR	+2.51%	+4.48%

References

[1] J. Buckley and I. James. Linear regression with censored data. *Biometrika*, 66(3):429–436, 1979.

[2] J. E. Fischer, C. Greenhalgh, and S. Benford. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, pages 181–190. ACM, 2011.

[3] I. R. James and P. Smith. Consistency results for linear regression with censored data. *The Annals of Statistics*, pages 590–600, 1984.

[4] N. Keiding, P. K. Andersen, and J. P. Klein. The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. *Statistics in medicine*, 16(2):215–224, 1997.

[5] A. Mehrotra, M. Musolesi, R. Hendley, and V. Pejovic. Designing content-driven intelligent notification mechanisms for mobile applications. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 813–824. ACM, 2015.

[6] A. Mehrotra, V. Pejovic, and M. Musolesi. Sensocial: a middleware for integrating online social networks and mobile sensing data streams. In *Proceedings of the 15th International Middleware Conference*, pages 205–216. ACM, 2014.

[7] M. Pielot, K. Church, and R. De Oliveira. An in-situ study of mobile phone notifications. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pages 233–242. ACM, 2014.

[8] L.-J. Wei. The accelerated failure time model: a useful alternative to the cox regression model in survival analysis. *Statistics in medicine*, 11(14-15):1871–1879, 1992.

[9] Y. Xu, M. Lin, H. Lu, G. Cardone, N. Lane, Z. Chen, A. Campbell, and T. Choudhury. Preference, context and communities: a multi-faceted approach to predicting smartphone app usage patterns. In *Proceedings of the 2013 International Symposium on Wearable Computers*, pages 69–76. ACM, 2013.