

January 8, 2017

Abstract

1

in recent studies, Bayesian learning has been used to analyze information aggregation in social networks, in which individuals modify their decision based on previous outcomes of other individuals in the network. (A cite is needed for the topic Bayesian Learning in Social Networks)

We introduce a novel method to model the churn rate based on Bayesian learning methods. Specifically, we view the churn as a decision making process: suppose an individual i decides whether to churn in a particular month t based on a random variable $u_{i,t} \in \{0, 1\}$, where $u_{i,t} = 0$ indicates a churn, otherwise a non-churn is indicated. $u_{i,t}$ is drawn as follows:

1. Assume two hyperparameters $\alpha_{i,t}$ and $\beta_{i,t}$;

2. Draw

$$p_{i,t} \sim \text{Beta}(\alpha_{i,t}, \beta_{i,t}) \quad (1)$$

where $\text{Beta}(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{B}(\alpha, \beta)}$, and $\text{B}(\alpha, \beta)$ is the normalization constant.

3. Draw

$$u_{i,t} \sim \text{Bernoulli}(p_{i,t}) \quad (2)$$

where $\text{Bernoulli}(x; p) = p^x(1-p)^{1-x}$ is the Bernoulli distribution.

The distribution of $u_{i,t}$ is a specification of the Beta-Binomial distribution, which has mean $\frac{\alpha}{\alpha + \beta}$ and variance $\frac{\alpha\beta}{(\alpha + \beta)^2}$. This distribution has good properties: on the one hand, the overall churn rate can be easily estimated using the hyperparameters of each individual; on the other hand, since the Beta distribution is the conjugate prior of the Bernoulli distribution, α and β can be conveniently updated given observed results. For simplicity, please refer to Pattern Recognition and Machine Learning for more details.

Updating the hyperparameters α and β can be viewed as a Bayesian learning process: an increase in α_i decreases i 's tendency to churn, while an increase in

β_i increases this tendency. Moreover, increasing α_i and β_i reduces the variance of the Beta distribution, indicating that the individual has a better estimation of p_i . We also take the network structure into account, where the outcome of nearer individuals have more influence.

The update process is described in Algorithm 1:

```

Data: The network  $G$ , and the churned individuals in the first month  $\Omega^{(t)}$ 
initialize  $t = 0$ ,  $\alpha_{i,0}$  and  $\beta_{i,0}$  for  $\forall i \in \Gamma(t)$ ;
while True do
    for  $\forall i \in \Gamma(t) \setminus \Omega^{(t)}$  do
         $\hat{\alpha} = \hat{\beta} = 0$ ;
        for  $\forall j \in \Gamma(t) \setminus \Omega^{(t)}$  do
             $\hat{\alpha} = \hat{\alpha} + 1/d_{ij}^{(\tau)^2}$ ;
        end
        for  $\forall j \in \Omega^{(t)}$  do
             $\hat{\beta} = \hat{\beta} + 1/d_{ij}^{(\tau)^2}$ ;
        end
         $\alpha_i = \alpha_i + \frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}}$ ;
         $\beta_i = \beta_i + \frac{\hat{\beta}}{\hat{\alpha} + \hat{\beta}}$ ;
    end
     $\Omega^{(t)} = \Phi$ ;
    for  $\forall i \in \Gamma(t)$  do
        Sample  $u_{i,t}$  according to 1 and 2;
        if  $u_{i,t} = 0$  then
             $i$  choose to churn;
             $\Omega^{(t+1)} = \Omega^{(t+1)} \cup \{i\}$ ;
        else
             $i$  choose not to churn;
        end
    end
     $t = t + 1$ ;
end

```

Algorithm 1: Algorithm for the Bayesian Churn Model