Contribution ID : **159**   Type : **Oral presentation**

Predicting Links in temporal Networks using a dynamic Stochastic Block Model

Longitudinal social network data, observed as discrete-time snapshots, can be modeled using the dynamic stochastic block model, which assumes that nodes are clustered into unobservable groups and that block memberships evolve according to a hidden Markov chain. This model can be used as an early warning system by predicting the occurrence of future links through the evolution of latent block memberships. Considering the sequence of binary adjacency matrices $\mathbf{Y}^{(t)}$, $t = 1, \ldots, T$, with entries $Y_{ij}^{(t)} = 1$ if nodes i and j are connected by an edge at time occasion t $(i, j = 1, \ldots, n)$, the model relies on node-specific discrete latent processes $U_i^{(1)}, \ldots, U_i^{(T)}$, whose parameters are the initial probabilities $\alpha_u = p(U_i^{(1)} = u)$, the transition probabilities $\pi_{u|v} = p(U_i^{(t)} = u | U_i^{(t-1)} = v)$, and the connection probabilities $\beta_{uv} = p(Y_{ij}^{(t)} = 1 | U_i^{(t)} = u, U_j^{(t)} = v)$. Estimation is carried out using an approximated maximum likelihood approach based on a variational approximation of the log-likelihood function.

We propose a novel link prediction method to forecast possible connections between nodes i and j at time $T+1$ by estimating the following edge probability
$$p_{ij}^{edge} = \beta_{\hat{u}_i^{(T+1)}, \hat{u}_j^{(T+1)}},$$
where $\hat{u}_i^{(T+1)} = \arg\max_{u=1,\ldots,k} \hat{q}_i^{(T+1)}(u)$ is the predicted block for node i at time occasion $T+1$ computed on the basis of the maximum a-posteriori rule. Here $\hat{q}_i^{(T+1)}(u)$ is the variational approximation of the conditional posterior probability that node i belongs to latent block u, predicted for time point $T+1$, given the observed network. A connection is then predicted when p_{ij}^{edge} exceeds a certain threshold, which can be determined based on F1 score.

The approach is validated through simulations across a wide range of scenarios and it is applied to public data about interactions among 164 ants observed over eleven days. Accurate predictions are obtained especially when the network is dense and relatively persistent.

Keywords/Topics

Link Prediction; Variational Expectation-Maximization Algorithm; Forecasting; Longitudinal Networks

Primary author(s) : BRUSA, Luca (University of Milano-Bicocca); PANDOLFI, Silvia (University of Perugia); PENNONI, Fulvia (University of Milano-Bicocca); BARTOLUCCI, Francesco (University of Perugia)

Presenter(s) : BRUSA, Luca (University of Milano-Bicocca)

Session Classification : Statistical approaches for clustering and community detection in complex networks

Track Classification : Statistical approaches for clustering and community detection in complex networks