Learning the Probability of Activation in the Presence of Latent Spreaders

When modeling the spread of a contagion among individuals in a community, each individual's probability of adopting the contagion depends on his innate susceptibility and his exposure to the contagion through his neighbors. While one often has knowledge regarding an individual's susceptibility, in many cases, whether or not an individual's neighbors are contagious is unknown or latent. We propose a new generative model in which we take into account the graph structure and model the contagious state of each neighbor in the graph as a latent variable. Using the neighbors' contagious state, we can estimate the individual's exposure to the contagion. Combined with an individual's characteristics, we estimate risk of infection as a function of both exposure and susceptibility. We propose a variational inference algorithm to learn the model parameters. Using simulated data, we show that our model outperforms benchmarks and is robust to variations in the relative importance of exposure and susceptibility in deciding the infection state. Applied to a real dataset of over 22,000 patients, we demonstrate the utility of our model for identifying 1) potential asymptomatic carriers of a hospital acquired infection, and 2) characteristics that are associated with an increased likelihood of being an undiagnosed source of contagion. Our results on real data highlight how our proposed method could be used to reduce the spread of infection within hospitals.