Inferring gang affiliation for violent events with incomplete data.

Data sets are often plagued with portions of missing or incomplete data. In this case, the data are assumed to be associated with one of N self-exciting point processes. This creates a network of processes such that each edge is associated with an independent self-exciting point process. The time and geographical location for all events are known, however the process affiliation is not known for some events. Previous work successfully used clustering in time resulting from the self-excitation to maximize the weights of the unknown events for each process. However, the authors assumed that the parameters of the process were known. Further, the maximization of the weights can be computationally costly. This work proposes a novel iterative method with a directly calculable score function assigning appropriate weights for process affiliation and approximating the crucial parameters needed to approximate the underlying point process. After testing this method on simulated data, it is applied to data on gang violence obtained from the LAPD. Gang violence perpetrated against rivals and the subsequent retaliations has been successfully described by self-exciting point processes and therefore fits ideally into this mathematical framework. (Received September 21, 2011)