How to detect information outbreaks in Twitter?

By Thibault Lasserre

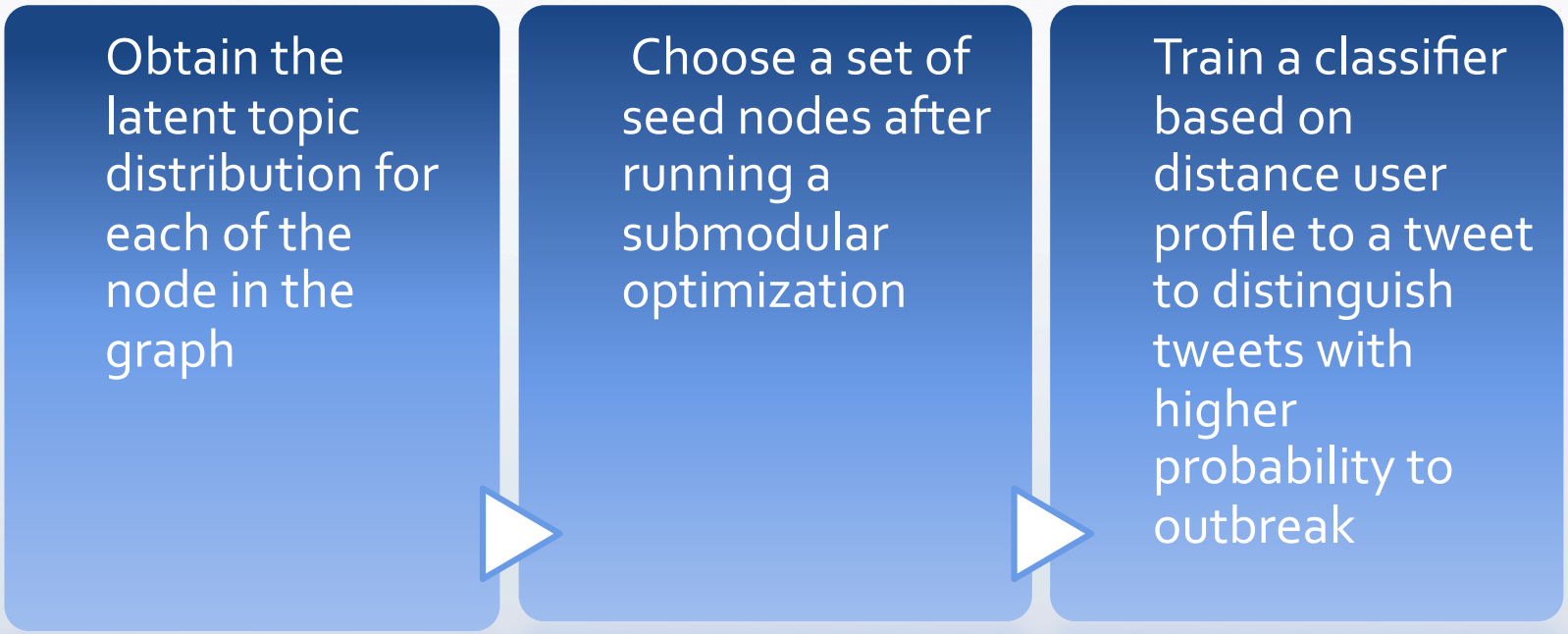
Problem Statement

- Everyone is receiving a tons of information on an unparalleled scale.
- The way to resolve the dilemma is to keep one's eyes on some trusted influential nodes in the information diffusion network.
- How to select a set of nodes to detect some dynamic spreading process over a network?

Hypothesis

- We can predict which tweet will be widely spread by focussing on influential nodes and their content-related information.

Method



Algorithms: Submodular optimization, Latent Dirichlet Allocation, SVM classifier, Linear Threshold and Independent Cascades

Data

- Tweet collection from Dr Aron Culotta
- Twitter Social Graph from Haewoon Kwak, et al.

Related Works

- Learning to Detect Information Outbreaks in Social Networks, Jiayuan Ma, Xincheng Zhang

- Maximizing the Spread of Influence through a Social Network, David Kempe, Jon Kleinberg, Eva Tardos

Timeline

- Week of Oct. 14th:Data Analysis
- Week of Oct. 21st: Compute the Latent Dirichlet Allocation and Analysis Submodular optimization
- Week of Oct. 28th: Compute the SVM Classifier
- Week of Nov. 4th: test and analyze the results
- Week of Nov. 11st: Write the final report