# Category-Level Transfer Learning from Knowledge Base to Microblog Stream for Accurate Event Detection

Weijing Huang, Tengjiao Wang, Wei Chen, Yazhou Wang

Peking University

DASFAA 2017

# Neural Network View of Topic Models (1/2)

The conditional probability $p(w|d)$ is the combination of **word-topic distribution** $p(w|t_i)$ and **topic-document distribution** $p(t_i|d)$.

$$p(w|d) = \sum_{i=1}^{K} p(w|t_i)p(t_i|d) \tag{1}$$

Equation (1) equals to the following form:

$$p(w|d) = \phi(w)\theta^T(d) \tag{2}$$

, where $\phi(w)$ is the row vector **word-topic distribution** $[p(w|t_1), \cdots, p(w|t_K)]$, and $\theta(d)$ is the row vector **topic-document distribution** $[p(t_1|d), \cdots, p(t_K|d)]$.

# Neural Network View of Topic Models (2/2)

In equation (2), the distributions $\phi(w)$ and $\theta(d)$ have the following explanation:

1. $\phi(w)$ functions as the look-up layer for words with the **sigmoid** activation functions.

2. $\theta(d)$ functions as the look-up layer for documents with the **softmax** activation function.

Symbols: