

# オンライン学習可能な多重スケールでの 時間発展を考慮したトピックモデル

岩田具治, 山田武士, 櫻井保志, 上田修功  
NTTコミュニケーション科学基礎研究所

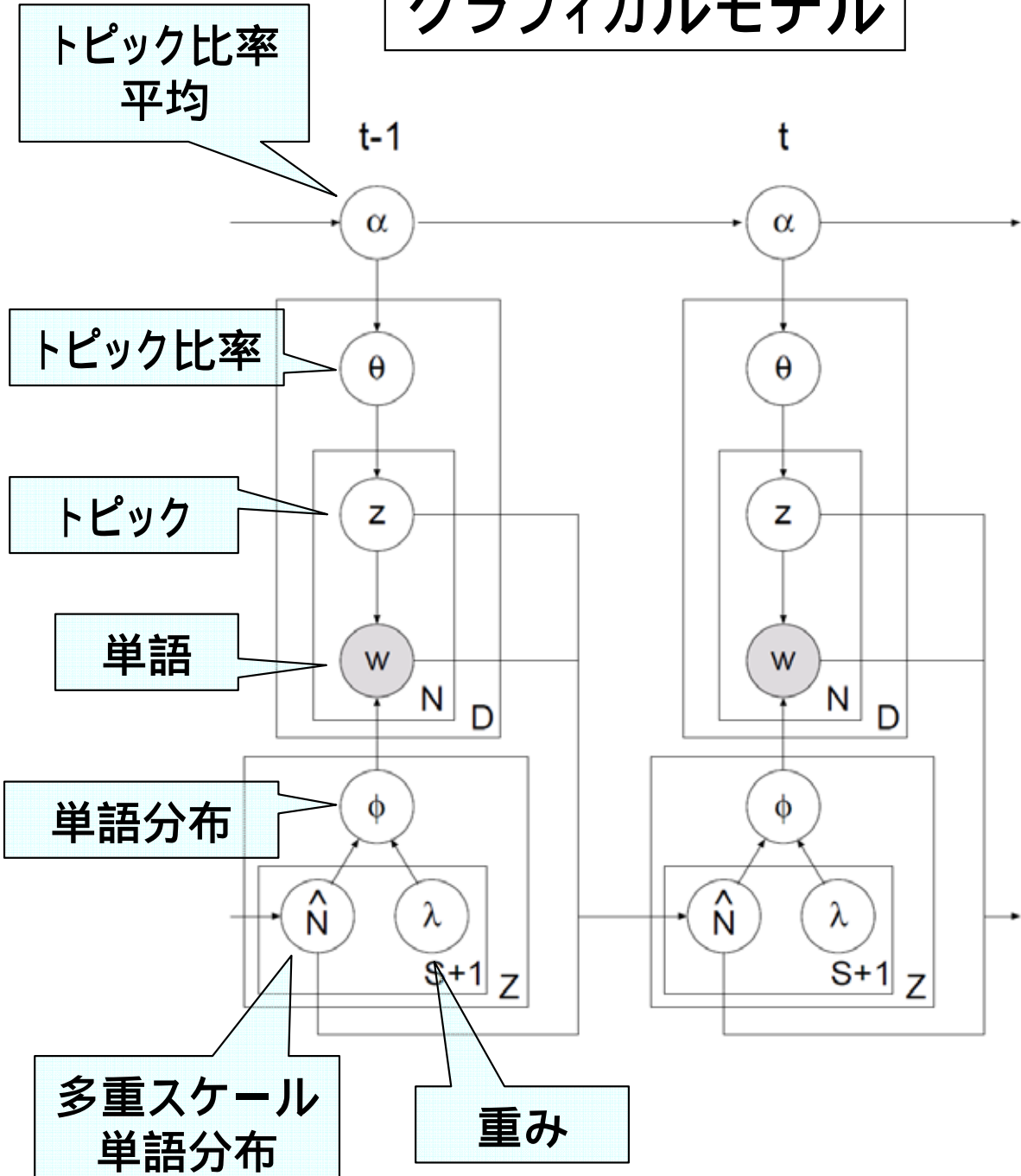
## 目的

- 日々蓄積される文書データを逐次的に解析

## 特徴

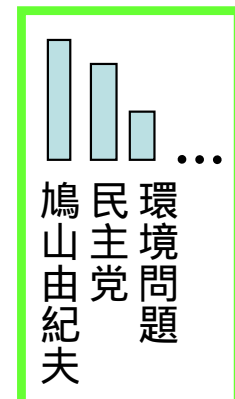
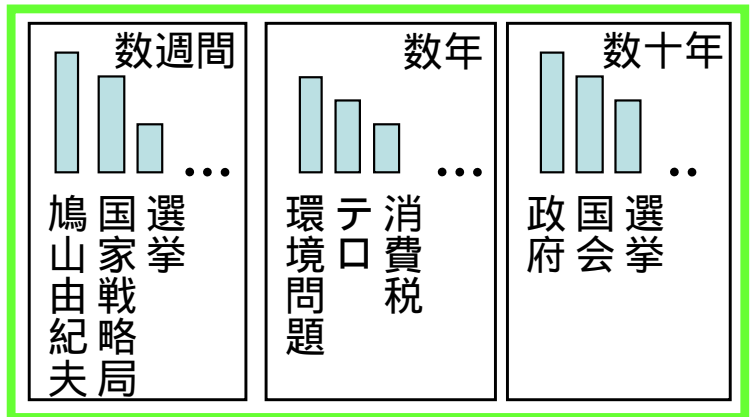
- 複数の時間スケールの時間発展を考慮することにより, **頑健なモデル推定**が可能
- オンライン学習により大規模ストリームデータでも**高速な学習**が可能

# グラフィカルモデル



事前単語確率は  
一時刻前の多重スケール  
単語確率の重み付き和を  
平均とするディリクレ分布

時刻t-1の複数の時間スケール  
での単語確率



例: 政治トピック

時刻tの単語確率