

## [论文名称] 連接社會媒體大資料中的“資料孤島”

[作者]: 桑基韜<sup>1</sup> 1 模式識別國家重點實驗室, 中科院自動化研究所

[摘要] 分析和利用社會媒體大資料的一個關鍵問題是“資訊孤島”: 每天有數以 EB 計的社交媒體資料生成並分散在上百種社交媒體服務和上萬個社交媒體網站上。綜合利用這些資料不僅要解決因數據異構性帶來的“語義鴻溝”, 更要解決因數據多源性帶來的“平臺鴻溝”。社交媒體大資料來自於個體的“小資料”, 同一個使用者往往同時參與到多個社交媒體網路的交互中。我們提出利用跨社會媒體網路的共同使用者作為連接“平臺鴻溝”橋樑, 採用類似群智感知的方法挖掘跨社交媒體網路的複雜知識關聯模式, 並應用於跨社會媒體網路協同的個性化資訊服務中。本報告將具體從以下兩方面進行介紹: (1) 跨網路關聯發現; (2) 跨網路使用者建模。