



微觀社群模擬器： 基於代理人模型之資訊擴散演化分析

Micro-Social Simulator:

Agent-Based Modeling of information Diffusion

學生：張璦云（F74122187）

指導教授：李政德 教授

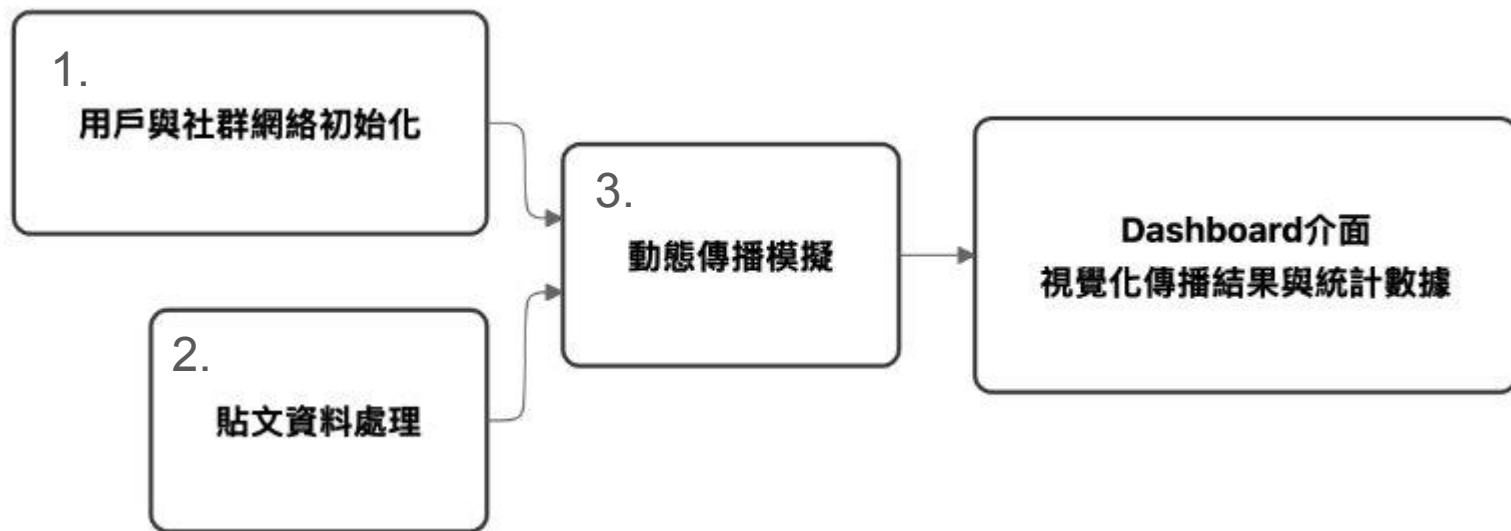
一、專題簡介

傳統傳播模型（如：IC model等）中，將資訊傳播過程簡化為機率事件

- 無法模擬用戶的異質性特質（如：年齡、職業、興趣等）
- 難以捕捉文章語意
- 無法處理多模態輸入
- 無法理解流行迷因

本專題的目標是解決這些痛點，建立起更貼近真實社群互動的動態傳播機率模型。

二、系統架構 (Workflow)



二、系統架構

1. 用戶與社群網絡初始化

此階段為模擬用戶行為與社交網絡環境的基礎設定。

- 手動輸入或批次建立用戶個人檔案(Profiles):
 - 年齡、性別、職業、教育程度、是否擁有車輛及感興趣的話題
 - 手動輸入: 輸入用戶自述, Qwen 2.5模型分析用戶潛在興趣話題

🔧 開始建立自訂虛擬用戶 Profile (寫入至 manual_test.json)

- ```
=====
```
1. 請輸入性別 (男性/女性): 女性
  2. 請輸入年齡 (如 25):  
-> 請輸入數字。
  2. 請輸入年齡 (如 25): 25
  3. 請選擇學歷 (選項: 國中及以下, 高中職, 專科, 大學, 研究所)  
輸入: 大學
  4. 請輸入職業 (例如: 軟體工程師, 學生...): 軟體工程師
  5. 請問是否有車? (y/n): n
- ```
=====
```

6. 請用一段話描述這個角色的個性、生活狀態或日常煩惱。

輸入描述: 平時喜歡觀看影視作品, 吃飯都會配youtube影片

```
{  
  "user_id": "UID_CUSTOM_7004",  
  "gender": "女性",  
  "age": 25,  
  "education": "大學",  
  "occupation": "軟體工程師",  
  "has_car": false,  
  "persona_description": "平時喜歡觀看影視作品, 吃飯都會配youtube影片",  
  "interests": [  
    "影劇與綜藝",  
    "短影音/TikTok",  
    "明星與網紅"  
  ]  
}
```

二、系統架構

1. 用戶與社群網絡初始化

- 基於用戶標籤 (Profile labels) 計算之用戶相似度 (U2U similarity)
- 指定隨機機率 (Random probability)

綜合以上兩者建立社群網絡結構。

二、系統架構

2. 貼文資料處理

本階段針對多模態的傳播資料進行特徵萃取與資訊整合。

❖ **文字處理：** 輸入資料的文字內容使用 Qwen 2.5 進行關鍵字提取。

。

❖ **圖片處理：**

- Paddle OCR 提取圖片中的文字，Qwen 2.5 進行雜訊清洗；
- CLIP 模型生成視覺嵌入向量（Vision Embedding）
- Intern VL2 產生圖片描述（Image Caption）。

二、系統架構

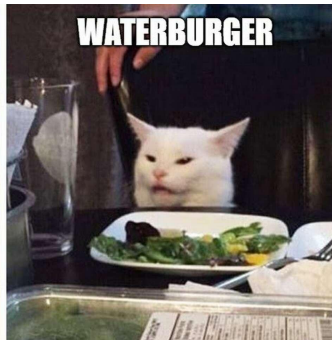
2. 貼文資料處理

❖ 梗圖辨識：

- Bge-m3 模型：清洗後的 OCR 文字與InternVL圖片描述轉換為文字嵌入向量（Text Embedding）。
- 將此Text embedding與前述的Vision embeddding，分別與 PostgreSQL 梗圖資料庫中現存的資料進行餘弦相似度（Cosine-similarity）計算，藉此判斷該圖片是否為一張梗圖（Meme）並分類其類型。



input
type: woman_cat
template
in DB



二、系統架構

2. 貼文資料處理

Q: Why not only vision embedding but **vision embedding + ocr + caption**?

A: 梗圖有二次創作文化



vis_emb cos-sim -> ocr text_emb+ caption text_emb : do cos-sim calculation

二、系統架構

3. 動態傳播模擬階段

基於前兩階段(social network/post features)，進行貼文擴散的時序迭代模擬。

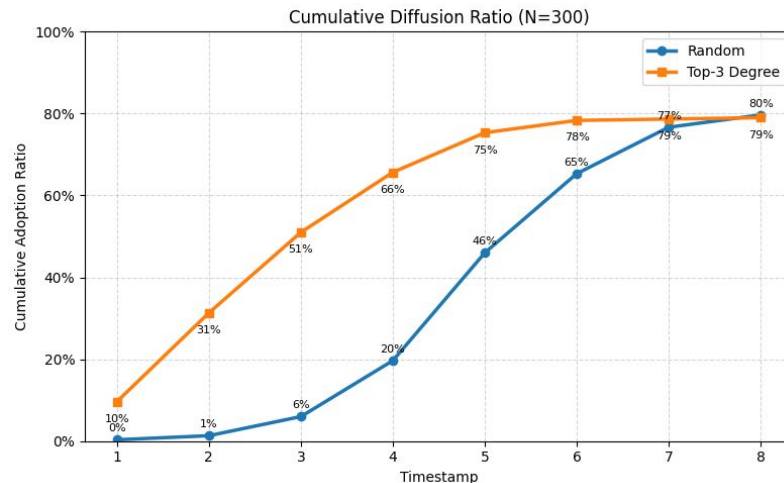
1. 在初始時間節點 ($T=0$) 指定起始活躍用戶節點 (Initial user nodes) 後，系統會隨著時間推移，逐一評估當前活躍節點的相鄰節點 (Neighbors)。(BFS)
2. 計算每位鄰居對該篇貼文的興趣程度或相似度 (U2P similarity)。
3. 用戶對貼文相似度 (U2P) 與用戶間相似度 (U2U) 綜合評估用戶接受傳播機率。
4. 若傳播成功，該相鄰節點將於下一時間點 ($T+1$) 被標記為新增的活躍節點。

此迭代評估迴圈將持續運作，直至當前時間點不再發生任何新的傳播行為，即宣告模擬終止。

三、Demo & 實驗結果展示

Expt. A. 資訊散播初始節點選

對同一社群、同篇貼文的反覆模擬初始節點數=3的傳播行為。



- 選擇高度中心性 (Degree)的初始節點較隨機選擇更快達到高覆蓋率的傳播階段，更快收斂。
- Timestamp = 3 時，前者與後者間的覆蓋率相差8.82倍，worst case中也至少有2倍以上的差距。

In real life -> KOL、網紅宣傳

三、Demo & 實驗結果展示

(鄰近用戶的同質性太高)

Expt. B & C 世代差異(Generational Gap)與同溫層效應(Echo Chamber Effect)

- 對一個符合台灣網路用戶年齡分佈的社群投放廢文、梗圖 [2025 網路用戶調查](#)
- 年長者即便被推送到貼文，但由於傳播的意願極低，資訊難以在年長者子社群擴散。
- 一個年齡結構高度兩極化（18-29 歲 或 60 歲以上）的社群網絡
- 散播高度符合年長者偏好的貼文（醫療養生）
- 顯著的結構性傳播差異：資訊於「年長者子網絡」中迅速且廣泛地蔓延；反觀左側的「年輕子網絡」，僅有極少數與右側相連的跨界節點接收到資訊，整體擴散動能微弱且進程顯著遲緩。

In real life -> 資訊繭房 (Information Cocoons) (演算法只推用戶喜歡的內容)

兩者皆會導致用戶個人視野限縮、加深確認偏誤。在雙方對立的事件中無法被推播到對立方的資訊內容



謝謝聆聽！

