

---

# 모바일 사용자를 위한 컨텍스트 기반 마이크로 블로그 토픽 검출 기법

## Context-based Microblog Hot Topic Detection for Mobile Users

한중현, Jonghyun Han\*, Xing Xie\*\*, 우운택, Woontack Woo\*\*\*

---

**요약** 최근 모바일 장치를 통한 마이크로 블로그 활용이 늘고 있지만, 모바일 장치가 지닌 하드웨어 제약으로 인해 여전히 모바일 브라우징에 어려움이 있다. 이를 해결하기 위해 모바일 사용자의 컨텍스트 정보를 활용하여 사용자의 관심 정보를 추론하는 연구가 활발히 진행되고 있다. 본 논문에서는 모바일 사용자의 컨텍스트를 이용하여 마이크로 블로그의 토픽을 추천하는 방법을 제안한다. 마이크로 블로그에서 사용자와 연관된 토픽을 추출하기 위해 제안한 방법은 사용자 위치, 행동, 기존에 작성한 블로그 그리고 사회적 관계 등의 사용자 컨텍스트를 모바일 장치로부터 얻어 활용한다. 모바일 장치로부터 얻어온 컨텍스트는 마이크로 블로그 검색 범위를 줄이는데 뿐만 아니라 사용자의 관심을 추론하는 경우에도 활용된다. 추론된 사용자의 선호도를 기반으로 검색된 결과의 우선순위를 다시 결정한다. 제안한 방법을 통해 모바일 사용자들은 사용자가 관심을 가질만한 토픽의 마이크로 블로그 정보를 얻을 수 있을 것으로 기대한다.

**Abstract** Mobile context-awareness becomes an important research topic since mobile information browsing is still difficult due to the limitations of mobile devices. On the other hand, it is easier to gather more user contexts because mobile devices are equipped with more sensors. In this paper, we introduce a real-time local hot topic browser which displays hot topics related to a mobile user. In order to retrieve user-related hot topics from microblogs, it exploits mobile user contexts such as location, activity, blogging history, and social relationship. Through taking advantage of these contexts, it reduces the search range of microblogs and infers user interests. It can filter out unrelated topics based on inferred preference. Through our proposed approach, a mobile user can be aware of topics related to interests surrounding the user.

**핵심어:** *Context-awareness, Microblog Topic Detection, Mobile Browser*

---

본 논문은 2010년 Microsoft Research Asia 인턴십 지원에 의하여 연구되었음.

\*주저자 : 광주과학기술원 정보통신공학과 U-VR 연구실 박사과정; e-mail: jhan@gist.ac.kr

\*\*공동저자 : Microsoft Research Asia Researcher; e-mail: xingx@microsoft.com

\*\*\*교신저자 : 광주과학기술원 정보통신공학과 U-VR 연구실 부교수; e-mail: wwoo@gist.ac.kr

## 1. 서론

스마트 폰의 확산으로 모바일 사용자들의 소셜 네트워크 서비스 이용이 늘어나고 있다. [1] 스마트 폰의 기술 발전으로 데스크톱 환경과 유사한 서비스를 사용자들은 제공받고 있지만, 모바일 장치가 가지는 작은 디스플레이 및 낮은 컴퓨팅 파워 등의 한계로 인해 많은 정보를 처리하고 확인하는데는 여전히 어려움이 있다. [2] 이러한 어려움을 극복하기 위해 사용자가 어떤 정보를 필요로 하는지 추론하여 적합한 소셜 콘텐츠를 제공해야 할 필요가 있다.

본 논문에서는 모바일 사용자를 위한 마이크로 블로그 토픽을 추출하는 방법을 제안한다. 제안한 방법은 모바일 폰으로부터 얻은 컨텍스트 정보를 이용하여 마이크로 블로그의 검색 범위를 축소시키고, 검출된 토픽들의 우선순위를 컨텍스트를 활용하여 결정하고 사용자에게 전달한다. 제안한 방법을 통해 모바일 사용자는 주변 환경에서의 일어나는 일에 대한 토픽을 쉽게 검색하여 확인할 수 있다.

## 2. 마이크로 블로그 토픽 검출 기법

우리는 모바일 폰으로부터 컨텍스트를 얻고, 얻은 컨텍스트들을 이용해 마이크로 블로그 토픽을 추출하고 모바일 폰에서 브라우징 할 수 있는 시스템을 디자인 및 개발하였다. 그림 1 은 구현된 컨텍스트 인지 마이크로 블로그 브라우저로 추출된 토픽과 함께 해당 마이크로 블로그를 지도 위에 보여준다.



그림 1. 컨텍스트 인지 마이크로 블로그 브라우저

### 2.1 지역 토픽 검출 방법

사용자가 위치한 지역의 마이크로 블로그 토픽을 추출하기 위하여 모바일 폰의 GPS, 전자 나침반 등을 이용한다. 모바일 폰으로부터 얻어온 사용자의 위치를 주소로 변환하고, 변환된 주소에서 포스팅 된 마이크로 블로그를 검색한다. 그리고 검출된 토픽의 지역성을 높이기 위하여 전역에서 포스팅 된 마이크로 블로그 역시 활용한다. 검출된 마이크로 블로그 토픽의 중요성을

확인하기 위해 토픽의 빈도수를 계산하고 [3], 계산한 결과를 전역 마이크로 블로그 빈도수 및 이전 시간 윈도우의 빈도수와 비교하여 중요도를 계산한다. 이 방법을 통해 지역성을 가지며 최근에 특별하기 이슈가 되는 토픽을 검출할 수 있게 된다.

### 2.2 컨텍스트 기반 토픽 선택 방법

추출된 토픽들 중 사용자가 관심을 가질만한 토픽을 선택하기 위해 우리는 컨텍스트 정보를 활용하여 토픽의 우선순위를 재결정한다. 이 때, 활용된 컨텍스트는 사용자의 이전 마이크로 블로그 히스토리, 사용자의 소셜 네트워크, 그리고 행동 정보이다. 각각의 컨텍스트에 해당하는 마이크로 블로그를 검색하고, 검색된 마이크로 블로그들의 단어 빈도수를 계산하여 검색된 마이크로 블로그들이 어떤 카테고리에 속하는지 확률을 계산한다. 이렇게 계산된 카테고리 벡터 정보는 사용자가 어떤 카테고리에 관심을 가지는지에 대한 정보를 가지게 된다. 이 정보와 검출된 토픽의 카테고리 정보의 유사도를 코사인 유사도 계산 방법을 이용하여 계산하고 이 값을 바탕으로 검출된 토픽의 우선순위를 다시 계산하게 된다. 재계산된 우선 순위를 바탕으로 상위 토픽들을 모바일 장치에 전달하여 사용자에게 제공한다.

## 3. 실험

제안된 방법의 검증을 위해 우리는 마이크로 블로그 데이터 집합을 75%의 트레이닝 집합과 25%의 테스트 집합으로 나누어 트레이닝 집합으로부터 마이크로 블로그 토픽을 검출하고 테스트 데이터 집합에 속한 각각의 마이크로 블로그와 검출된 토픽간의 유사도를 계산하여 제공한 토픽의 적합성을 확인하였다. 그림 2 는 제안한 방법에 대한 실험 결과를 보여주고 있다. 그림 2 를 통해 볼 수 있듯이 컨텍스트 정보의 활용을 통해서 검출된 토픽들의 관련성을 개선시킬 수 있는 것을 확인하였다.

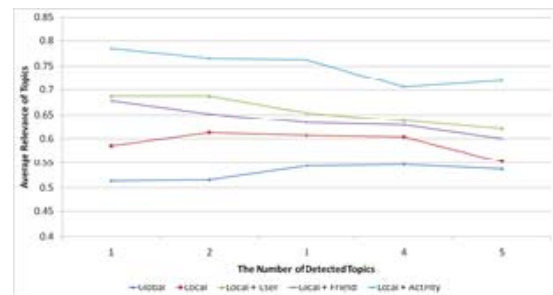


그림 2. 검출된 토픽의 유사도 실험 결과

#### 4. 결론

본 논문에서는 모바일 사용자의 컨텍스트를 활용한 마이크로 블로그 토픽 추출 기법을 제안하였다. 제안된 기법은 사용자의 위치 정보를 통해 마이크로 블로그 검색 범위를 줄이고, 이전 마이크로 블로그 히스토리, 소셜 네트워크, 행동 정보를 이용하여 검출된 토픽들의 우선순위를 사용자 관심도에 따라 재계산하였다. 제안된 기법을 이용하여 검출한 마이크로 블로그 토픽들의 적합성을 확인하는 것을 통해 제안된 방법의 유용성을 검증하였다. 추후 연구로는 사용자 스타디를 통해 제안된 기법의 유용성 실험이 필요하다.

#### 참고문헌

[1] A. Java, X. Song, T. Finin, and B. Tseng. "Why we twitter: understanding microblogging usage and

communities". In Proceedings of WebKDD/SNA-KDD 2007, pages 56-65, New York, 2007.

[2] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton. "Cyberguide: a mobile context-aware tour guide". *Wirel. Netw.*, 3(5):421-433, 1997.

[3] K.-Y. Chen, L. Luesukprasert, and S.-c. T. Chou. "Hot topic extraction based on timeline analysis and multidimensional sentence modeling". *IEEE Trans. on Knowl. and Data Eng.*, 19(8):1016-1025, 2007.