

## 지리적 공간과 장치 정보를 사용한 개선된 트윗 봇 검출

이알찬<sup>1</sup> · 서고은<sup>1</sup> · 신원용<sup>2\*</sup> · 김동건<sup>3</sup> · 조재희<sup>4</sup>

### Improved Tweet Bot Detection Using Geo-Location and Device Information

Al-Chan Lee<sup>1</sup> · Go-Eun Seo<sup>1</sup> · Won-Yong Shin<sup>2\*</sup> · Donggeon Kim<sup>3</sup> · Jaehee Cho<sup>4</sup>

<sup>1</sup>Department of Mobile Systems Engineering, Dankook University, Yongin 448-701, Korea

<sup>2\*</sup>Department of Computer Science and Engineering, Dankook University, Yongin 448-701, Korea

<sup>3</sup>Department of Statistics and Information Science, Dongduk Women's University, Seoul 136-714, Korea

<sup>4</sup>Department of Management, Kwangwoon University, Seoul 139-701, Korea

#### 요 약

온라인 소셜 네트워크 서비스 중 하나인 트위터는 가장 보편적으로 사용되는 마이크로 블로그인데, 트위터의 개방적 구조로 인해 자동화 프로그램인 트윗 봇이 많이 생성되고 있다. 이 트윗 봇은 적법한 봇과 악성 봇으로 분류되는데, 이 중 악성 봇은 일반 사용자들에게 많은 양의 스팸 정보나 유해한 콘텐츠를 배포하기 때문에 트윗 봇을 검출하는 작업은 반드시 필요하다. 기존 연구에서는 시간적 정보를 활용하여 사람과 트윗 봇을 분류하였다. 본 논문에서는 먼저 사용자들의 고 정밀 위치 정보를 알려주는 공간 태그된 트윗 정보를 활용하여 트위터 사용자들의 정확한 위치를 알아낸다. 그리고, 각 사용자의 공간 변수에 대한 엔트로피 값 및 사용자의 장치 정보를 사용하여 새로운 봇 검출 알고리즘을 제안한다. 주요 결과로써, 시간 정보만을 이용한 기존 연구결과보다 각 신뢰도별 봇 검출 확률 및 거짓 경보 확률이 모두 우수하게 나타난다.

#### ABSTRACT

Twitter, one of online social network services, is one of the most popular micro-blogs, which generates a large number of automated programs, known as tweet bots because of the open structure of Twitter. While these tweet bots are categorized to legitimate bots and malicious bots, it is important to detect tweet bots since malicious bots spread spam and malicious contents to human users. In the conventional work, temporal information was utilized for the classification of human and bot. In this paper, by utilizing geo-tagged tweets that provide high-precision location information of users, we first identify both Twitter users' exact location. Then, we propose a new tweet bot detection algorithm by using both an entropy based on geographic variable of each user and device information of each user. As a main result, the proposed algorithm shows superior bot detection and false alarm probabilities over the conventional result which only uses temporal information.

**키워드** : 장치 정보, 공간 정보, 공간 태그된 트윗, 온라인 소셜 네트워크, 트위터, 트윗 봇

**Key word** : Device information, geographic information, geo-tagged tweet, online social network, Twitter, tweet bot

Received 10 August 2015, Revised 01 September 2015, Accepted 14 September 2015

\* Corresponding Author Won-Yong Shin(E-mail:wysin@dankook.ac.kr, Tel:+82-31-8005-3253)

Department of Computer Science and Engineering, Dankook University, Yongin 448-701, Korea

Open Access <http://dx.doi.org/10.6109/jkiice.2015.19.12.2878>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.  
Copyright © The Korea Institute of Information and Communication Engineering.

## I. 서론

온라인 소셜 네트워크 (OSN: Online Social Network)는 인터넷 상의 상호 의사 소통 및 정보 교류에 있어 유용한 사회기반 시스템이다. 이러한 OSN은 컴퓨터, 모바일 장치, 그리고 태블릿과 같은 인터넷이 가능한 장치 사용의 급증으로 인해 더욱 대중화 되었다. OSN 사용자들은 그들의 프로필과 콘텐츠를 게시하고 사람들과 링크를 공유하는 등의 활동을 하며, 이러한 활동을 통하여 사용자들은 사회적 관계를 유지하게 된다[1]. OSN은 사용자들의 활동을 기반으로 수집된 빅 데이터를 가지고 있는데, 이는 여러 가지 분석에 유용하게 사용된다. OSN 중에서 “트위터 (Twitter)”는 가장 보편적으로 사용되는 마이크로 블로그 중 하나인데, 사용자들이 트윗 (tweet) 이라는 140자 이내의 단문메시지를 전송함으로써 개인의 견해나 정보를 공유하고 다른 사용자들을 팔로우 (follow) 함으로써 그들의 트윗을 받을 수 있다[2]. 트위터에서는 다른 OSN과 달리 사용자 개인의 요구에 의해 타 사용자와의 일방적인 관계 형성이 가능하기 때문에 유동적인 정보의 흐름이 가능하다.

인간, 컴퓨터 바이러스, 교통, 인구 예측 등과 같은 다양한 응용이 가능하기 때문에 인간 개개인의 이동성에 대한 많은 연구가 진행되어 왔다. 인간 이동성에 대한 이전 연구들에서는 모바일 폰을 활용한 추적 기술[3, 4]을 사용하는 것에 초점을 맞추었다. 이전 연구들에서는 분석에 필요한 데이터가 부족하거나 사용자의 정확한 위치 정보를 얻지 못하는 등의 결점이 있었다. 반면, 공간 태그된 트위터(geo-tagged Twitter)[5-8]는 사용자의 동의 하에 스마트폰의 GPS interface를 통해 10미터 이내의 정확한 공간 정보 (위도 및 경도)를 제공해주며, 거대한 샘플과 함께 포괄적인 메타데이터를 제공한다. 공간 태그된 트윗은 전체 트윗 양의 대략 1% 정도를 차지하지만, 스마트 기기 및 모바일 앱의 폭발적인 보급으로 인해 이제는 분석에 유효한 충분한 양의 데이터 확보가 가능해졌다.

### 1.1. 관련 연구

기존 연구에서는 트위터 계정의 특징을 분석하고 그 계정들을 사람, 트윗 봇(tweet bot), 사이보그(cyborg)로 분류하는 방법을 제안하였다[9]. 위의 세 가지 집단을 구분 짓는데 필요한 세 가지 요소는 다음과 같다. 첫 번째

요소인 엔트로피 (entropy)인데, 엔트로피란 정보에 대한 불규칙성을 나타내는 것으로 불규칙적일수록 엔트로피 값은 증가하게 된다. 트윗 간 시간 (tweet interval time) 변수에 대한 엔트로피 값을 계산하여 사람과 봇의 시간적 패턴, 즉 주기성을 비교하였다. 두 번째 요소는 스팸 검출 요소로, 이를 통해 트윗의 텍스트가 스팸 콘텐츠를 가지고 있는지 여부를 판단한다. 세 번째 요소는 계정 속성 요소인데, URL 등의 포함 여부를 조사한다. 위 세 가지 요소의 각 계정별 데이터 중 자신을 팔로우하는 사용자의 수와 자신이 팔로우 하는 사용자의 수의 비율 (Follower-Friend Ratio)을 사용하여 봇 검출을 수행하였다.

### 1.2. 주요 제안 사항

본 논문에서는 공간 태그된 트윗을 통해 도출한 지리적 공간 정보와 트윗 생성 시 사용된 장치 분류를 기반으로 하는 새로운 봇 검출 알고리즘을 소개한다. 트윗 봇은 트윗 간 이동거리가 0에 가깝거나 (즉, 이동하지 않고 동일 지점에서만 트윗을 보내거나) 사람에 비해 거대 스케일로 비교적 규칙적인 이동을 하는 경향이 있다. 반면 사람은 걷거나 교통수단을 이용하여 이동하기 때문에 트윗 간 거리의 분산이 높게 나타난다. 또한, 데이터 셋 중 사용자가 어떤 장치로 트윗을 생성하였는지를 표기해주는 source 필드를 통해 사용자별 장치를 분석할 수 있다. 사람이 사용하는 장치 분포도에서 0.5% 이상의 확률 값을 가지는 장치들만을 선별하여 이 장치들을 사용하는 사용자는 사람이라고 가정할 수 있다. 이러한 두 가지 특성을 봇 검출 알고리즘에 적용하고자 한다.

먼저, 구해진 트윗 간 거리를 사용하여 불확실성 정도를 나타내는 엔트로피 (entropy)를 계산한다. 훈련 과정에서는 지정한 신뢰도가 가능하도록 하는 거리 변수에 대한 엔트로피 값을 나타내는 임계값 (threshold)을 설정한다. 테스트 과정에서는 두 가지 정보를 사용하는데, 앞서 설정한 임계값보다 엔트로피 값이 작은 레코드를 가지면서 앞서 언급한 사람이 주로 사용하는 몇 가지 장치 집합에 포함하지 않는 사용자들을 봇으로 분류한다.

주요 결과로써, 각 신뢰도 별 거짓 오류 확률 (FAP: False Alarm Probability)을 보인다. 또한 봇 검출 확률 (DP: Detection Probability)과 FAP 사이 상관관계를 보인다. 여기에서 FAP란 사람을 봇으로 잘못 검출할 확률이다. 봇 DP가 높다는 것은 더 정확한 트윗봇 검출이

가능하다는 것이고, FAP 가 높을수록 사람이 봇으로 오인될 확률이 높은 것을 의미한다. 기존 연구 [9]에서의 결과와 비교를 수행함으로써 제안한 알고리즘의 우수성을 검증한다.

본 논문의 구성은 다음과 같다. II장에서는 데이터셋에 대해 소개한다. III장에서는 새로운 봇 검출 알고리즘을 포함한 연구방법론을 제안한다. IV장에서는 분석결과를 정리하고, V장에서는 본 논문을 요약 및 마무리한다.

## II. 데이터셋

트위터 분석을 위한 데이터 수집 프로세스로 어플리케이션의 접근 횟수에 대해 제한이 없고 실시간 이벤트 검출에 용이한 장점이 있는 트위터 Streaming API (Application Programming Interface)를 이용하여 트위터 네트워크 크롤링을 진행하였다. 크롤러는 트위터 API로부터 R의 StreamR 패키지를 이용하여 일정시간 단위로 JSON (JavaScript Object Notation) 파일을 저장한다. 이후 수집된 JSON 파일에서 필요한 필드를 parsing하여 하나의 분석용 표로 완성하고, 분석에 필요한 데이터마트를 구성한다. 분석 시에는 알고리즘 구현을 위해 R 소프트웨어를 사용한다.

본 연구에서는 2014년 9월 5일부터 11일까지 (대략 일주일) 트위터 사용자로부터 기록된 거대한 양의 공간 태그된 트윗으로 구성된 데이터셋을 사용하였다. 일주일 가량의 단기간 데이터셋은 사람과 트윗 봇의 통계적 특성을 분석하기에 충분하다. 수집된 각 사용자 데이터 중 트윗이 10번 이상 발생한 사용자의 정보만을 이용하여 데이터셋을 구성하였다.<sup>1)</sup> 각 사용자별 트윗 간 거리에 따른 엔트로피를 계산하는데 있어 트윗 레코드가 적으면 엔트로피 값에 대한 정확성이 떨어지기 때문이다. 데이터 수집은 영국 London 내 트위터 사용자들을 대상으로 진행하였다. 데이터셋에서 각 트윗 기록은 트윗이 보내진 장소 및 시간을 가진다. 이 정보에 기반하여 각 사용자의 위치 정보를  $L = (x_{ij}, y_{ij})$ 와 같이 나타내었다. 여기에서,  $x_{ij}$ 와  $y_{ij}$ 는 각각 사용자  $i$ 의  $j$ 번째 트윗이 발생되었을 때의  $x$ -좌표 및  $y$ -좌표이고, 이 위치

정보는 위도와 경도를 나타낸다. 데이터셋은 사용자 트윗에 대한 많은 정보를 포함하고 있으며, 트윗의 메타 데이터에서 본 연구에 필요한 필드 값을 아래와 같이 추출한다.

- *user\_id\_str*: 사용자 ID
- *lat*: 사용자의 위도
- *lon*: 사용자의 경도
- *source*: 트윗을 생성한 장치

이제 기존 연구 [9]와 마찬가지로 ground-truth 방법을 통해 사람 셋과 봇 셋을 분류하는 과정을 설명한다. 트위터 페이지에 게시하는 트윗 콘텐츠를 분석하여 규칙적으로 같은 메시지나 URL을 반복적으로 트윗하는 사용자를 트윗 봇으로 분류하였다. 또한, 트윗의 텍스트가 스팸 콘텐츠를 포함하고 있는지를 판단하여 트윗 봇으로 간주하였다. 사람 셋은 다시 두 개로 분할하는데, 첫 번째 절반의 집합은 훈련 셋 (training set)으로, 나머지 절반의 집합은 전체 봇 셋과 함께 테스트 셋 (test set)으로 구성하였다. 사람 셋 및 봇 셋으로 구분한 이후 사람으로 보이지만 트윗 간 속도가 300km/h보다 빠른 사용자는 사람일 가능성이 낮다고 판단하여 봇으로 재분류하였다. 기준이 300km/h인 이유는 그 이상의 비이상적인 속도를 보인다면 사람이 이용 가능한 교통수단이 없다고 판단하였기 때문이다. 본 논문에서는 분류 기준에 의하여 전체 1,007개의 사용자 계정 중 사람 계정 892개와 트윗 봇 계정 115개로 데이터셋을 구성하였다. 따라서 훈련 셋에는 사람 계정 수의 절반인 446개가 포함되고, 테스트 셋에는 나머지 446개의 사람 계정과 115개의 봇 계정을 합한 561개의 계정이 포함된다. 사람 셋의 절반을 나누는 기준은 사용자 ID를 기준으로 무작위로 분류하였다.

## III. 연구 방법론

본 장에서는 지리적 공간 정보 및 장치 정보를 이용하여 트윗 봇을 검출하는 전체적인 과정을 설명한다.

### 3.1. 엔트로피 계산 과정

먼저 공간에서의 사람과 트윗 봇의 트윗 패턴을 관찰

1) 10번이란 횟수는 임의적으로 분석하기에 충분한 숫자로 판단하여 정한 것이다.

및 분석해 보도록 한다. 트윗 봇은 트윗 간 이동거리가 0에 가깝거나 사람에 비해 거대 스케일로 비교적 규칙적인 이동을 하는 경향이 있다. 따라서 봇의 트윗 간 거리 변수에 대한 엔트로피도 사람보다 훨씬 작다고 판단할 수 있다.

**Table. 1** Partition of each distance section

$D_{ij}$	Section interval (km)	Number of sections
$0 < D_{ij} < 0.1$	0	1
$0.1 \leq D_{ij} < 1$	0.1	9
$1 \leq D_{ij} < 3$	0.5	4
$3 \leq D_{ij} < 10$	1	7
$10 \leq D_{ij} < 800$	10	79

**Table. 2** Entropy of each user according to the dataset

Dataset	Definition
Training set	$E^{TR} = \{e_1^{TR}, \dots, e_{U_1}^{TR}\}$
Test set	$E^{TE} = \{e_1^{TE}, \dots, e_{U_2}^{TE}\}$

사용자  $i$ 가  $j$ 번째 트윗을 보냈을 때의 위치 정보가  $(x_{ij}, y_{ij})$ 이고  $(j+1)$ 번째 트윗을 보냈을 때의 위치 정보가  $(x_{i(j+1)}, y_{i(j+1)})$ 라 할 때, 두 지점 사이의 지리적 거리  $D_{ij}$ 는 구면 코사인 법칙 (spherical law of cosines)을 사용하여 구할 수 있다. 거리  $D_{ij}$ 는 최대 800km로 제한하였으며  $D_{ij}=0$ 인 경우를 먼저 분할한 후 표 1에서 보이는 것과 같이 구간별로 총 101 구간으로 차등 분할하였다.

이제 데이터셋 (훈련 셋, 테스트 셋)에 따른 사용자별 엔트로피를 표 2에서 정의한다. 구체적으로, 훈련 셋 및 테스트 셋에서의 엔트로피 집합을 각각  $E^{TR}$ 과  $E^{TE}$ 로 나타내자. 훈련 셋에서  $i$ 번째 사용자의 트윗 간 거리 변수에 대한 엔트로피 값을  $e_i^{TR}$ 로 표기한다. 마찬가지로 테스트 셋에서  $i$ 번째 사용자의 거리에 대응하는 엔트로피 값을  $e_i^{TE}$ 로 표기한다. 또한,  $U_1$ 과  $U_2$ 는 각각 훈련 셋과 테스트 셋에 포함된 전체 사용자 수를 의미한다. 앞서 II장에서 언급한바와 같이  $U_1 = 446$ ,  $U_2 = 561$ 로 주어진다.

다음은 불확실성을 나타내는 엔트로피를 구하는 기

본적인 공식이다. 훈련 셋 내 레코드를 사용하여  $i$ 번째 사용자의 거리에 대응하는 엔트로피는 다음과 같이 주어지게 된다.

$$e_i^{TR} = - \sum_{j=1}^n P_{D_{ij}} \log P_{D_{ij}} \quad (1)$$

여기서  $P_{D_{ij}}$ 는  $j$ 번째 인덱스에 대한 변수  $D_{ij}$ 의 확률 분포,  $n$ 은 데이터 총 개수를 나타낸다. 표 3에서는 각 사용자별 엔트로피를 구하는 방법을 보여준다.

**Table. 3** Entropy computing algorithm for each user

```

Input:  $P_{D_{ij}}, i \in \{1, \dots, U_1\}, j \in \{1, \dots, n\}$ 
Output:  $e_i^{TR}$ 
Initialization:  $i \leftarrow 1, j \leftarrow 1, e_{di}^{TR} \leftarrow 0$ 
00: for  $i \leftarrow 1$  to  $U_1$  do
01: for  $j \leftarrow 1$  to  $n$  do
02:  $e_i^{TR} \leftarrow e_i^{TR} - P_{D_{ij}} \log P_{D_{ij}}$ 
03: end for
04: end for
    
```

**Table. 4** Training process of detection algorithm

```

Input:  $\{e_1^{TR}, \dots, e_{U_1}^{TR}\}$ 
Output: d_entropy
Initialization: prob  $\leftarrow 1, h\_count \leftarrow 0, i \leftarrow 1$ 
00: while (prob > A/100)
01: for  $i \leftarrow 1$  to  $U_1$  do
02: if  $e_i^{TR} > d\_entropy$  and  $i == \text{'Human'}$ 
03: h_count++
04: end if
05: end for
06: prob  $\leftarrow h\_count / U_1$ 
07: d_entropy  $\leftarrow d\_entropy(1 + \Delta)$ 
08: end while
    
```

### 3.2. 트윗 봇 검출 과정

본 절에서는 앞서 계산한 각 사용자별 엔트로피 및 장치 정보를 이용하여 트윗 봇 검출하는 새로운 알고리즘에 대해 설명한다.

• 훈련 과정: 먼저 거리 변수  $D_{ij}$ 를 사용한 훈련 과정을 수행하여 각 신뢰도별 임계값을 구한다. 임계값이란 지정한 신뢰도가 가능하도록 하는 엔트로피 값이며, 신뢰도란 사람 사용자가 지정한 범위 내에 포함될 확률이며

신뢰도가 높을수록 지정한 범위 내의 사람 사용자가 많음을 뜻한다. 예를 들어, 신뢰도가 80%가 되도록 하는 변수  $D_{ij}$ 에 대한 임계값은 훈련 셋에서 엔트로피  $e_i^{TR}$ 의 최댓값으로부터 사람의 80%가 포함되도록 하는 그러한 엔트로피의 최솟값으로 설정된다. 이 때 최댓값을 기준으로 신뢰도를 지정한 이유는 봇이 사람에 비해 보다 비교적 정지 상태에서 트윗을 하는 경향 때문이다. 표 4에서는 신뢰도가  $A\%$ 인 경우에 대한 훈련 과정을 보인다.

**Table. 5** Test process of detection algorithm

<b>Input:</b> $\{e_1^{TE}, \dots, e_{U_2}^{TE}\}$
<b>Output:</b> $P_{bd}, P_{fa}$
<b>Initialization:</b> $d\_entropy \leftarrow \eta, h\_count \leftarrow 0, b\_count \leftarrow 0$
<b>00:</b> for $i \leftarrow 1$ to $U_2$ do
<b>01:</b> if $E_i^{TE} < d\_entropy$ and $i$ 's device $\notin DV$ and $i == 'Bot'$
<b>02:</b> $b\_count++$
<b>03:</b> else if $E_i^{TE} < d\_entropy$ and $i$ 's device $\notin DV$ and $i == 'Human'$
<b>04:</b> $h\_count++$
<b>05:</b> end if
<b>06:</b> end for
<b>07:</b> $P_{bd} \leftarrow b\_count / U_{2b}, P_{fa} \leftarrow h\_count / U_{2h}$

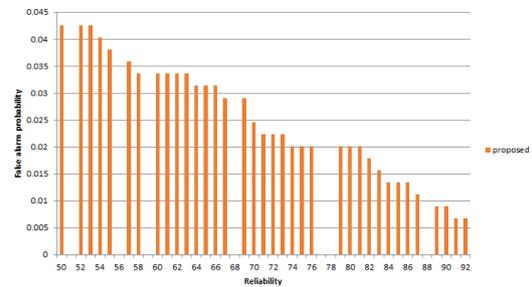
• 테스트 과정: 다음으로 앞서 설정한 임계값과 선별된 장치 집합을 바탕으로 검출 알고리즘을 수행한다. 테스트 과정을 설명하기 전에 먼저 *source* 필드를 통한 사용자 구분 기준을 설명하도록 한다. 사람이 사용하는 장치 분포도에서 0.5% 이상의 확률 값을 가지는 장치들만을 선별하여 이 장치들을 사용하는 사용자는 사람이라고 가정하였으며, 선별된 사람의 장치 집합을 *DV*로 정의하였다.

$$DV = \left\{ \begin{array}{l} \text{Twitter for iPhone, Twitter for Android,} \\ \text{Instagram, Twitter for ipad, Foursquare,} \\ \text{Twitter for Android Tablets,} \\ \text{Twitter for Windows phone, iOS} \end{array} \right\}$$

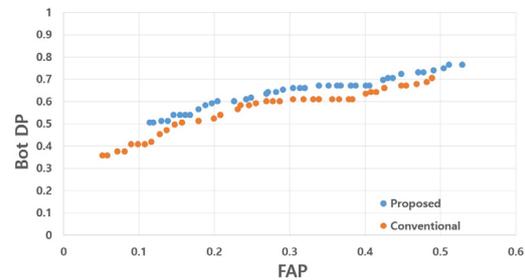
이제 본 논문에서 제안하는 거리 변수의 엔트로피 및 장치 집합 *DV*에 따라 검출을 진행한다. 앞서 설명했듯이 트윗 간 거리 변수의 분산이 작을수록 트윗 봇일 확률이 높으므로 앞서 설정한 엔트로피 임계값보다 작은 사용자는 봇일 가능성이 높다. 또한, 사용자가 트윗 시

사용한 장치가 집합 *DV*에 포함되지 않을 때 봇일 가능성이 높다. 따라서 임계값보다 작은 엔트로피를 가지면서 장치가 *DV*에 속하지 않을 때 봇으로 판별한다.

트윗 봇인 경우  $b\_count$ 를 증가시켜 봇 DP  $P_{bd}$ 를 구하고 사람이 오인되었을 경우  $h\_count$ 를 증가하여 FAP  $P_{fa}$ 를 구하게 된다. 거리 변수에 대한 임계값을  $\eta$ 로 설정한 경우의 검출 알고리즘은 표5에서 자세히 보여진다. 이 때,  $U_{2h}$ 와  $U_{2b}$ 는 테스트 셋에 존재하는 사람 수와 트윗 봇 수를 나타낸다.



**Fig. 1** FAP for each reliability



**Fig. 2** FAP vs. Bot DP

#### IV. 분석결과

본 장에서는 [9]에서의 사람과 봇의 시간적 패턴을 분석한 결과와 제안하는 알고리즘의 개선된 성능 결과를 보인다. 구체적으로 각 신뢰도별 봇 DP 및 봇 DP와 FAP 간 상관관계를 보인다. [9]에서는 봇 검출을 위해 시간 정보 이외에도 트윗 내 스팸 콘텐츠 유무, URL 등의 계정 속성을 이용하였으나, 이러한 두 가지 요소는 제안하는 알고리즘에도 함께 적용되어 추가적인 성능

개선이 가능하다. 하지만 본 장에서는 새롭게 소개된 공간 및 장치 정보를 사용하였을 때의 성능에 초점을 맞추고자 한다.

#### 4.1. 신뢰도별 FAP

그림 1은 지리적 공간 및 장치 정보를 이용하여 제안한 알고리즘에 대해 신뢰도별 FAP 측면에서 분석한 결과를 보여준다. 신뢰도가 높을수록 사람이 사람으로 인식될 확률이 높아지고, 신뢰도가 낮아질수록 불안정한 봇 검출이 되므로 FAP 또한 증가하게 된다. 또한, FAP의 최댓값은 0.04로 매우 작은 값을 확인할 수 있다.

#### 4.2. 봇 DP와 FAP 사이의 상관관계

그림 2에서는 봇 DP와 FAP 간 상관관계를 보여준다. 특정 임계값보다 엔트로피 값이 작은 사용자를 트윗 봇으로 판별하는 본 논문의 알고리즘에서는 FAP가 높아질수록 그만큼 봇 검출 될 확률도 높아지게 됨을 알 수 있다. 또한, 기존 기술 [9]와 제안한 알고리즘을 비교해 보면 제안한 알고리즘이 같은 FAP에서 더 높은 봇 DP를 보인다. 따라서 사람이 봇으로 오인 될 확률이 같을 때 더 높은 봇 DP를 보여 더 안정적이고 정확한 검출이 가능하다는 것을 확인할 수 있다.

## V. 결 론

본 논문에서는 트윗 봇 검출을 위한 새로운 알고리즘을 제안하였다. 트윗 간 시간 정보만을 이용한 기존 연구 [9]에서 시도되지 않은 트윗 간 거리 정보 및 트윗을 생성한 장치 정보를 사용하였는데, 거리 변수에 대한 엔트로피 값을 계산 후 엔트로피 값과 선별된 장치 집합을 통해 봇 검출을 수행하였다. 주요 결과로써, 기존 결과 대비 각 신뢰도별 FAP가 크게 향상된 것을 확인할 수 있었고 같은 FAP에서 더 높은 봇 DP를 보임을 검증하였다.

본 논문에서 사용된 사람 셋과 봇 셋 분류는 트윗 컨텐츠를 일일이 수작업 표시로 수행하였기에 데이터셋 크기가 충분하지 않았다. 좀 더 신뢰성 있는 결과를 위해 더욱 방대한 셋 분류 작업이 필요한데, 향후 연구로 크라우드 소싱 (crowdsourcing)을 사용한 셋 분류를 수행할 예정이다.

## ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2014R1A1A2054577) and by the Ministry of Science, ICT & Future Planning (MSIP) (2015R1A2A1A15054248).

## REFERENCES

- [1] C. Wilson, B. Boe, A. Sala, K. P. N. Puttaswamy, and B. Y. Zhao, "User interaction in social networks and their implication," in *Proceedings of the 4th ACM European Conference on Computer Systems (EuroSys '09)*, Nuremberg, Germany, pp. 205-218, Mar./Apr. 2009.
- [2] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?," in *Proceedings of the 19th International World Wide Web Conference (WWW2010)*, Raleigh, NC USA, pp. 591-600, Apr. 2010.
- [3] M. C. Gonzalez, C. A. Hidalgo, and A. L. Batabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, pp. 591-600, Apr. 2010.
- [4] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi, "Human mobility, social ties, and link prediction," in *Proceedings of the 17th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD2011)*, San Diego, CA USA, pp. 1100-1108, Aug. 2011.
- [5] B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti, "Geo-located Twitter as proxy for global mobility patterns," *Cartography and Geographic Information Science*, vol. 41, no. 3, pp. 260-271, May 2014.
- [6] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth, "Understanding human mobility from Twitter," *PLOS ONE*, vol. 10, no. 7, pp. 1-16, July 2015.
- [7] W.-Y. Shin, B. C. Singh, J. Cho, and A. M. Everett, "A new understanding of friendships in space: Complex networks meet Twitter," *Journal of Information Science*, vol. 41, no. 6, pp. 751-564, Dec. 2015.
- [8] S. Y. Jeon, A. C. Lee, G. E. Seo, and W. Y. Shin, "Relationship between tweet frequency and user velocity on Twitter," *Journal of the Korea Institute of Information and*

*Communication Engineering*, vol. 19, no. 6, pp. 1380-1386, Jun. 2015.

[9] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no.6, pp. 811-824, Dec. 2012.



**이알찬(Ai-Chan Lee)**

2012년~현재 단국대학교 국제학부 모바일시스템공학전공 학사과정  
※관심분야 : 통신이론, 빅데이터 분석



**서고은(Go-Eun Seo)**

2012년~현재 단국대학교 국제학부 모바일시스템공학전공 학사과정  
※관심분야 : 통신이론, 빅데이터 분석



**신원용(Won-Yong Shin)**

2002년 연세대학교 기계전자공학부 학사  
2004년 KAIST 전자전산학과 석사  
2008년 KAIST 전자전산학부 박사  
2008년 2월~4월 Harvard University 방문연구원  
2008년 9월~2009년 2월 KAIST BK 정보전자연구소 박사후연구원  
2009년 3월~4월 KAIST 고성능집적시스템연구센터 선임급 위촉연구원  
2009년 5월~2011년 10월 Harvard University Postdoctoral Fellow  
2011년 10월~2012년 2월 Harvard University Research Associate  
2012년 3월~현재 단국대학교 국제학부 모바일시스템공학전공/대학원 컴퓨터학과 조교수  
※관심분야 : 정보이론, 통신이론, 신호처리, 모바일 컴퓨팅, 빅데이터 분석



**김동건(Donggeon Kim)**

1986년 연세대학교 경영학과 학사  
1990년 Virginia Tech 통계학과 석사  
1995년 Virginia Tech 통계학과 박사  
1996년~현재 동덕여자대학교 정보통계학과 교수  
※관심분야 : 통계계산, 데이터마이닝, 빅데이터 분석



**조재희(Jaehee Cho)**

1983년 연세대학교 경영대학 학사  
1985년 Miami University (Ohio) 경영학 석사  
1991년 University of Nebraska-Lincoln 경영정보학 박사  
1994년~현재 광운대학교 경영대학 교수  
2001년 UC Berkeley 교환교수  
2011년~2013년 한국데이터마이닝학회 학회장  
※관심분야 : Twitter data analytics, 데이터웨어하우스, 데이터 시각화, 네트워크 분석