

안철민, 김영남

본문에서는 입력된 어휘를 자모단위로 분리하고 표준화한 후 류사도를 평가하는 방법으로 변형된 금지어를 검출하는 한가지 방법을 제안하였다.

사용자가 《빠꾸기》라는 단어가 금지어사전에 등록되었다는것을 알고 《빠구기》

로 변형하여 전송하였다고 하자.

이 자모분리단계를 거친 후 어휘는 《배 나 나 나 나》의 형태로 된다.

걸음 2 자모를 표준형으로 변환한다.

자모를 표준형으로 변환한다는것은 발음상 유사한 자음 혹은 모음들을 하나의 모임으로 보고 매 모임에 대표자모를 설정하며 모임에 속한 자음 혹은 모음을 이 대표자모로 치환하는것을 말한다.

조선어자모의 대표자모에 의한 표준화변환규칙은 표 1과 같다.

표 1. 조선어자모의 대표자모에 의한 표준화변환규칙

자 음								
ㄱ	1.0	ㄱ	ㄲ	1.0	ㄲ	ㄴ	1.0	ㄴ
ㄴ	0.9		ㄷ	0.9		ㄷ	0.9	
ㄷ	0.8		ㄸ	0.8		ㄹ	0.8	
ㅈ	1.0	ㅈ						
ㅉ	0.9							
ㅊ	0.8							
모 음								
ㅏ	1.0	ㅏ	ㅑ	1.0	ㅑ	ㅓ	1.0	ㅓ
ㅓ	0.8		ㅕ	0.8		ㅗ	0.8	
			ㅛ	0.7		ㅜ		
ㅗ	1.0	ㅗ	ㅛ	1.0	ㅛ			
ㅛ	0.8		ㅕ	0.8		ㅓ		
ㅕ	0.7		ㅑ	0.7		ㅓ		
			ㅓ	0.6				
받 침								
ㄱ	1.0	ㄱ	ㄲ	1.0	ㄲ			
ㄴ	0.9		ㄷ	0.9				
ㄷ	0.8		ㄴ	0.7				
			ㅈ	0.7				
			ㅉ	0.7				
			ㅊ	0.7				
			ㅋ	0.7				
			ㅇ	0.7				

표 1에서는 조선어의 특성에 기초하여 서로 련관되는 자모들을 모임으로 설정한 다

음 매 모임마다 대표자모를 정하고 매 자모와 대표자모와의 류사도를 결정하였다.

류사도결정방법은 다음과 같다.

① 자음에서 순한 소리자음에 대응한 된소리자음에 0.9, 거센소리자음에 0.8을 설정한다.

② 모음에서 밝은 소리모음에 대응한 어두운 소리모음에 0.8을 설정한다.

③ 모음에서 밝은 소리모음과 류사한 겹모음에 0.7을 설정한다.

④ 받침에서 자음변환규칙을 그대로 주는것을 원칙으로 하면서 《ㄷ》받침으로 발음되는 모든 자음들을 하나로 묶어준다.

표 1에서 제기한 류사도값들은 그자체 수값으로서는 큰 의의를 가지지 않는다. 왜냐하면 이 값은 조선어자모체계에서 《ㄱ》과 《ㄲ》을 사람이 어느 정도로 비슷하게 인식하는가 하는 주관적인 느낌과 연결되기때문이다. 더우기 이 값은 알고리즘의 마지막단계에서 리용하는 턱값과 연결된다. 여기서는 0.5~0.9의 범위에서 류사도를 설정하였다.

이 단계에서 우의 레문은 표준형 《ㄷ ㄱ ㄱ ㄱ ㄱ》로 변환된다.

변환된 표준형을 금지어사전에서 검색한다.

변형어휘《ㄷ ㄱ ㄱ ㄱ ㄱ》과 사전어휘《뽀꾸기》는 표준형으로 변환하면 《ㄷ ㄱ ㄱ ㄱ ㄱ》과 일치하므로 어휘는 금지어후보로 선택된다.

류사한 금지어휘와 후보어휘사이에도 그 문자렬(자모)길이가 다른 경우에는 이 단계를 그대로 통과하게 되는것이다.

실례로 《뽀꾸기》를 변형하여 《ㄷ ㄷ ㄱ ㄱ ㄱ ㄱ》로 전송하는 경우 표준형 《ㄷ ㄱ ㄱ ㄱ ㄱ》 혹은 《ㄷ ㄷ ㄱ ㄱ ㄱ ㄱ》과 일치하지 않으며 따라서 금지어후보로 선택되지 않는다.

이를 위하여 이 단계에서 표준형문자렬을 금지어사전에서 보통의 문자렬검색방식으로 찾는것이 아니라 류사도비교에 의한 검색방식을 리용할수 있다.

일정한 턱값을 설정하고 두 문자렬사이의 류사도가 턱값을 넘는 경우 해당 어휘를 금지어후보로 선택하는 방식이다.

그러나 이 방식은 실천적으로 체계자원을 대단히 많이 소모하는것으로 하여 대규모본문봉사체제에 적합치 않으며 금지어검출요구수준이 최대한으로 높은 체제에서 사용할수 있다.

어휘와 표준형과의 류사도는 다음과 같이 계산된다.

$$R(f, b) = \frac{\sum_{i=1}^m r(f_i, b_i)}{m}$$

여기서  $b$ 는 금지어휘,  $b_i$ 는 금지어의  $i$ 번째 자모,  $f$ 는 후보어휘,  $f_i$ 는 후보어휘의  $i$ 번째 자모,  $m$ 은 어휘의 길이이다.

$r(f_i, b_i)$ 는  $f_i$ 와  $b_i$ 의 류사도로서 다음의 공식에 의해 계산된다.

$$r(f_i, b_i) = \begin{cases} 1.0, & f_i = b_i \\ r(f_i, d_i) \times r(b_i, d_i), & f_i \neq b_i \end{cases}$$

여기서  $d_i$ 는 후보어휘의  $i$ 번째 자모의 표준형이다. 이 식의 의미는 문자들이 일치하는 경우 1.0, 일치하지 않는 경우에는 각각 표준형  $d_i$ 와의 류사도를 계산하여 그 적을 취한것

이다.

우의 레문을 놓고보자. 어휘 《배 ㅏ ㅓ ㅓ ㅓ ㅓ》와 사전어휘 《배 꾸기》와의 류사도는 다음과 같이 계산된다.

$$r=(0.9(\text{배}, \text{배})+1.0(\text{ㅏ}, \text{ㅏ})+1.0(\text{ㅓ}, \text{ㅓ})+1.0(\text{ㅓ}, \text{ㅓ})+1.0(\text{ㅓ}, \text{ㅓ})+1.0(\text{ㅓ}, \text{ㅓ}))/6=0.983$$

계산된 류사도와 톡값을 비교하여 금지어여부를 결정한다.

계산된 류사도가 톡값을 넘는 경우 이 어휘는 해당한 금지어에 대한 변형어휘일 가능성이 높으므로 금지어로 판단한다.

금지어로 판단한 후에 본문전송보류, 관리자에게 통지 등의 후속조치는 해당 본문통신봉사체계의 특성에 따라 선택할수 있다. 여기서 톡값은 0.9로 설정하였다.

실례로 어휘 《배 ㅏ ㅓ ㅓ ㅓ ㅓ》는 류사도가 톡값 0.9를 넘는것으로 하여 금지어로 판단된다.

### 3. 실험결과 및 분석

#### 1) 실험환경구성

특정한 과학분야(체신)의 어휘 2 500개를 선택하여 어휘자료기지로 모의하였다.(표 2)

어휘자료기지에 속하는 단어 200개를 선택하여 금지어로 설정하고 그중에서 100개를 여러가지 형식으로 변형하여 자료기지화하였다.(표 3)

표 2. 모의된 어휘자료기지

번호	용어
1	전자주사
2	무계계수기
3	에쓰디엘
⋮	⋮
2 498	송수신절환기
2 499	거리영상
2 500	회선다중화장치

표 3. 금지어자료기지(변형금지어포함)

번호	금지어	변형금지어
1	1차속도	1차속도
2	2진그린함수	2진그린함수
3	교환회로망	교환회로망
⋮	⋮	⋮
101	먼거리통화	먼거리 ㅓ 통화
⋮	⋮	⋮
199	제한기	제한 ㅓ ㅓ
200	닭알겹질	닭아르겹질

해당 분야의 문장 500개를 선택하여 통보문모임으로 설정하였다.

금지어검출체계성능시험에 사용된 자료기지특성은 표 4와 같다.

표 4. 금지어검출체계성능시험에 사용된 자료기지특성

특성지표	값
금지어개수	200
통보문개수	2 500
통보문 금지어 총개수	288
변형된 금지어개수	100
통보문 평균어휘개수	12.1

## 2) 실험방법

자동본문전송체계가 통보문을 금지어검출 체계봉사기에 발송하면 봉사기는 금지어검출을 진행한 후 그 결과를 기록한다.

금지어검출결과는 표 5와 같다.

시험결과 변형된 금지어 100개를 포함한 288개의 금지어중 281개가 검출되었다. 원본그대로의 금지어 188개는 전부 검출되었으며 변형된 금지어 100개중 93개가 검출되었다.

또한 금지어와 유사한 1개의 정상어휘가 금지어로 검출되었다.

표 5. 금지어검출결과

구 분	값
금지어 개수	288
검출된 금지어 개수	281
변형금지어 개수	100
검출된 변형금지어 개수	93
정상단어가 금지어로 검출된 비율/%	0.34
금지어 검출률/%	97.57
금지어 검출오율률/%	1.39
변형금지어 검출률/%	93
변형금지어 검출오율률/%	1

## 맺 는 말

제안된 금지어검출체계는 조선어통보문의 어휘변형행위에 대하여 금지어검출정확성이 높다.

또한 초당 1 000개까지의 통보문을 처리할수 있는 성능을 갖춘것으로 하여 수만~수십만가입자규모의 통신봉사체계들에서 안정하게 동작할수 있다.

## 참 고 문 헌

- [1] G. Navarro, C. Edgar; Theor. Comput. Sci., 352, 1, 266, 2006.
- [2] C. Lai; Know. Based Syst., 20, 3, 249, 2007.
- [3] A. C. Kigerl; Criminal Justice Policy Review, 27, 8, 791, 2016.

주체108(2019)년 2월 5일 원고접수

## A Method of Detecting Modified Vocabulary by Similarity Valuation in Text Communication System

An Chol Min, Kim Yong Nam

We valued the similarity between vocabularies on Korean alphabet and proposed an algorithm which detected the modified vocabulary based on a swear dictionary, and valued effectiveness by examination.

Key words: swear, modified vocabulary, similarity