

## 한국어 역사 소설에서 공간적 배경 인식 기법\*

김서희\*\* · 김승훈\*\*\*

### A Recognition Method for Korean Spatial Background in Historical Novels\*

Seo-Hee Kim\*\* · Seung-Hoon Kim\*\*\*

#### ■ Abstract ■

Background in a novel is most important elements with characters and events, and means time, place and situation that characters appeared. Among the background, spatial background can help conveys topic of a novel. So, it may be helpful for choosing a novel that readers want to read. In this paper, we are targeting Korean historical novels. In case of English text, It can be recognize spatial background easily because it use upper and lower case and words used with the spatial information such as Bank, University and City. But, in case Korean text, it is difficult to recognize that spatial background because there is few information about usage of letter. In the previous studies, they use machine learning or dictionaries and rules to recognize about spatial information in text such as news and text messages. In this paper, we build a nation dictionaries that refer to information such as 'Korean history' and 'Google maps.' We Also propose a method for recognizing spatial background based on patterns of postposition in Korean sentences comparing to previous works. We are grasp using of postposition with spatial background because Korean characteristics. And we propose a method based on result of morpheme analyze and frequency in a novel text for raising accuracy about recognizing spatial background. The recognized spatial background can help readers to grasp the atmosphere of a novel and to understand the events and atmosphere through recognition of the spatial background of the scene that characters appeared.

Keyword : Data Mining, Korean Historical Novels, Korean Linguistic Feature, Spatial Background

Submitted : February 1, 2016

1<sup>st</sup> Revision : March 6, 2016

Accepted : March 10, 2016

\* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2015년도 문화 콘텐츠산업기술지원사업(자유공모)의 연구결과로 수행되었음.

\*\* 단국대학교 대학원 컴퓨터학과 석사과정

\*\*\* 단국대학교 응용컴퓨터공학과 교수, 교신저자

## 1. 서론

소설을 구성하는 가장 중요한 요소는 인물, 사건, 배경이다. 그 중 배경은 소설에 등장하는 인물이 생각하고 행동하는 시간과 장소 또는 등장인물이 처해있는 상황을 의미하며, 시간적 배경, 공간적 배경, 사회적 배경, 자연적 배경 등으로 분류할 수 있다. 소설의 배경은 인물의 행동과 소설 속 사건을 실감나게 전달하는 역할을 하며, 사건 전개를 암시하는 역할을 한다. 또한 비유적인 표현이나 상징적으로 표현되기도 하며, 주제를 나타내고 소설의 분위기를 형성하는 데에도 영향을 미친다. 그 중 공간적 배경은 나라, 지방, 지형적 특징 등 사건이 일어나는 자연 환경이나 생활환경을 의미하는 것으로 소설의 주제를 드러내는 데 도움을 준다.

기존의 외국어 문서를 대상으로 한 연구에서는 통계적 기법을 이용하여 기계학습을 통해 공간적 정보를 인식했다. 영어의 경우 대소문자를 구분하며, Bank, University, City 등 공간적 정보와 함께 나타나는 단어들이 있기 때문에 공간적 정보의 인식이 비교적 쉬울 수 있으나, 한국어의 경우 문자의 사용에 대한 정보가 부족하기 때문에 공간적 정보의 인식이 어렵다. 한국어 문서를 대상으로 한 기존 연구들에서는 사전과 규칙 등을 이용하여 공간적 정보를 인식했다(Jang et al., 2005; Lee et al., 2000). 그러나 이러한 연구들에서는 신문 기사를 분석 대상으로 하여 현재 실존하는 공간을 인식하기 때문에 소설 속 과거 지명이나 실존하지 않는 허구의 지명은 인식하지 못한다.

따라서 본 논문에서는 한국어로 된 소설의 본문을 대상으로 하여, 사전을 이용한 기법과 조사 패턴을 이용한 기법을 결합한 공간적 배경 인식 기법을 제안한다. 한국어의 특성에 따라 공간적 배경에 함께 사용되는 조사를 파악하고 이를 통하여 공간적 배경 후보를 인식한다. 또한 정확도 향상을 위해 형태소 분석 결과와 소설 텍스트에서 공간적 배경 후보의 빈도수를 이용하여 공간적 배경을 인식한다. 소설의 여러 장르 중 공간적 배경이

특히 중요한 역할을 하는 것은 역사 소설이기 때문에 본 연구는 역사소설을 대상으로 한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 기존의 문서에서 공간적 정보를 인식하는 기법들에 관한 관련 연구들에 대해 나열한다. 제 3장에서는 본 논문에서 제안하는 한국어 역사 소설에서 공간적 배경 인식 기법에 대해 설명한다. 제 4장에서는 제안하는 기법을 이용한 실험 절차를 기술하고 결과를 확인한다. 마지막으로 제 5장에서는 본 논문의 결론 및 향후 연구 방향에 대해 제시한다.

## 2. 관련 연구

문서에서 공간적 정보를 인식하는 기존 연구는 크게 사전과 규칙을 이용하는 규칙 기반 방식과 학습 모델을 이용하는 통계적 기법으로 나눌 수 있다(Jang et al., 2005).

### 2.1 규칙 기반 방식을 통한 공간적 정보 인식

규칙 기반 방식이란 고유명사 사전이나 결합 사전 등을 이용하여 규칙을 수작업으로 구축하여 공간적 정보를 인식하는 방식을 말한다.

기존 연구에서는 공간적 정보인 지명을 인식하기 위해 우편번호부에 등록되어 있는 지역을 대상으로 사전을 구축했다. 품사가 태깅된 어절 단위의 신문 기사를 입력으로 하여 고유명사나 보통명사를 찾고, 미리 구축된 사전으로 검색했다. 사전에 검색된 결과가 없을 경우, 지명 후보 단어의 앞, 뒤에 나타나는 단어를 문맥규칙에 적용시켰다. 적용된 문맥규칙은 앞 단어의 품사가 동사이거나 뒷 단어의 품사가 명사인 경우 지명으로 인식하는 것을 말한다. 예를 들어, ‘성살이라 부르는 오션산 밑의’라는 문장에서 ‘오션산’이 지명 후보일 때, ‘부르는’이 동사이므로 지명으로 인식하는 것이다(Jang et al., 2005). 또 다른 기존 연구에서는 지명 사전과 국가명 사전, 지명결합명사 사전, 지명명사 사전, 용언의 하위 범주화 사전 등을 이용하여 한국

어 신문 기사에서 공간적 정보를 인식했다. 품사가 태깅된 어절 단위의 문장을 입력으로 하여 고유명사나 보통명사를 찾고, 사전 검색 및 용언의 사용 규칙, 개체명 간의 관계 등 사람이 직접 기술한 규칙을 적용하여 공간적 정보로 인식했다(Lee et al., 2000).

## 2.2 통계적 기법을 통한 공간적 정보 인식

통계적 기법이란 학습 말뭉치로부터 공간적 정보에 대한 지식을 학습한 뒤 모델을 이용하여 공간적 정보를 인식하는 기법을 말한다.

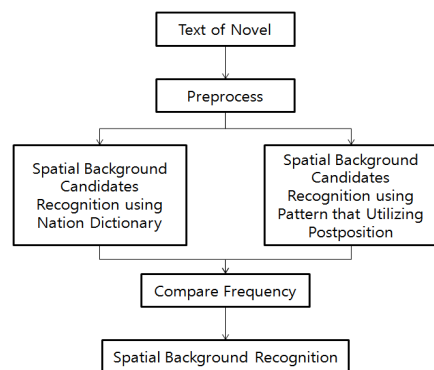
통계적 기법을 사용한 기존 연구에서는 HMM (Hidden Markov Models), MaxEnt(Maximum Entropy Based Models), SVM(Support Vector machines), ME(Maximum Entropy Classifier), NB(Naive Bayes), CRF(Conditional Random Field based models) 등을 사용하여 공간적 정보를 인식했다. 기존 연구에서는 품사가 태깅된 단어를 입력으로 하여 MUC-7(Seventh Message Understanding Conference) 말뭉치와 HMM의 2가지 변형 모델을 사용하여 공간적 정보를 인식했다(Todorovic et al., 2008). 또한 입력된 타밀어 문서를 형태소 분석하고, 2가지 규칙을 적용한 뒤, HMM을 이용하여 공간적 정보 인식하는 연구도 있었다(Srinivasagan et al., 2014). 또 다른 기존 연구에서는 각각의 HMM, MaxEnt, SVM, CRF 모델과 규칙을 결합하여 영어와 힌디어에서의 공간적 정보를 인식하고 그 결과를 비교했다(Prasad et al., 2015). 이와 유사하게 SVM, ME, NB 모델을 결합하여 힌디어에서 공간적 정보를 인식한 기존 연구도 있었다(Jaber and Saad, 2016). 학습된 말뭉치로부터 맥락적 특성을 추출하고, MaxEnt로 특성을 학습한 뒤, 학습된 특성과 기존에 구축된 사전, 그리고 간단한 규칙을 기반으로 중국어 문장에서의 공간적 정보를 인식한 연구도 있었으며(Zhang and Zhang, 2007), 실존하는 한국의 지명으로 구축된 사전과 CRF 모델을 적용하여 한국어

로 된 휴대전화 단문 메시지를 대상으로 공간적 정보를 인식하는 연구도 있었다(Kim and Kim, 2011).

그러나 저자에 따라 변형될 수 있고 각각 다른 문체적 특징을 가질 수 있는 한국어 소설을 대상으로는 하는 기존 연구는 찾을 수 없었다. 따라서 본 논문에서는 한국어 역사 소설을 대상으로, 사전 및 조사 패턴을 이용한 공간적 배경 인식 기법에 대해 제안한다.

## 3. 한국어 역사 소설에서의 공간적 배경 인식 기법

본 논문에서 제안하는 한국어 역사 소설에서의 공간적 배경 인식 기법은 국가명 사전을 이용한 기법과 조사 패턴을 이용한 기법으로 구성되어 있으며, 전체 흐름은 <Figure 1>과 같다. 한국 역사 소설의 텍스트가 입력되면 전처리 과정을 거친 뒤, 국가명 사전을 이용한 기법과 조사 패턴을 이용한 기법에 적용한다. 각 기법을 통해 인식된 공간적 배경 후보의 빈도수를 기반으로 한국어 역사 소설의 공간적 배경을 인식한다.



<Figure 1> Spatial Background Recognition

### 3.1 공간적 배경 인식을 위한 소설 텍스트 전처리

소설 텍스트에 사용되는 문장은 출판사, 저자에 따라 서로 다른 문장 부호를 사용하며, 문체에 따라

문장의 길이가 달라질 수 있다. 따라서 본 논문에서는 한국 역사 소설에서 공간적 배경을 인식하기 전, 소설 텍스트에 사용된 모든 문장을 동일한 형식으로 일치시키기 위한 전처리를 한다. 평문과 대화문을 동일한 형식으로 일치시키기 위해 ,(마침표),?(물음표),!(느낌표),“(큰따옴표), ‘(작은따옴표)를 기준으로 문장을 분리한다. 또한 문장을 단순화시키기 위해 ,(쉼표)를 기준으로 모든 문장을 분리한다.

### 3.2 국가명 사전을 이용한 공간적 배경 인식 기법

국가명 사전은 한국 역사 국가명 사전과 세계 국가명 사전으로 분류된다. 한국 역사 국가명 사전과 세계 국가명 사전은 ‘한국의 역사’, ‘Google 지도’ 등 여러 자료를 참고하여 구성된 것으로, <Table 1>과 <Table 2>는 그 예시이다.

<Table 1> Examples of Korean Historical Nation Dictionary

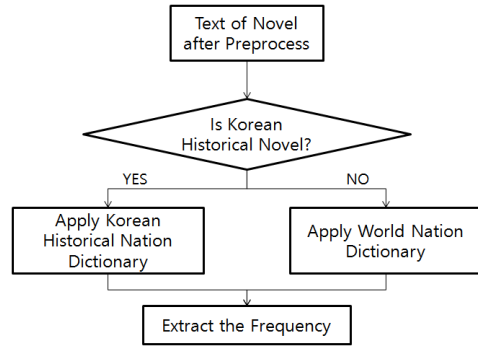
Dangun Josun (단군조선)	Wiman Josun (위만조선)
Jinguk(진국)	Buyeo(부여)
Goguryeo(고구려)	Baekje(백제)
Silla(신라)	...

<Table 2> Examples of World Nation Dictionary

Korea(대한민국)	Philippines(필리핀)
India(인도)	Papua New Guinea (파푸아뉴기니)
Australia(오스트레일리아)	New Zealand(뉴질랜드)
Canada(캐나다)	...

국가명 사전을 이용한 공간적 배경 인식 기법은 <Figure 2>과 같다. 전처리 된 소설 텍스트가 입력되면, 해당 소설이 한국 소설인지 아닌지에 따라 각각 다른 국가명 사전을 적용한다. 입력된 소설이 한국 소설인 경우 한국 역사 국가명 사전을

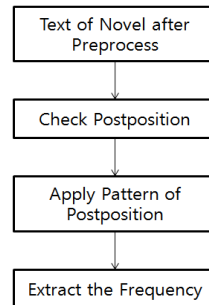
적용하고, 외국 소설인 경우 세계 국가명 사전을 적용하여 공간적 배경 후보를 인식한다. 또한, 인식된 공간적 배경 후보가 소설 텍스트에서 얼마나 많이 나타나는지 빈도수를 추출한다.



<Figure 2> Spatial Background Candidates Recognition using Nation Dictionary

### 3.3 조사 패턴을 이용한 공간적 배경 인식 기법

한국어의 경우 영어와 달리 대소문자의 구분이 없으며, Bank, City 등과 같은 공간을 나타내는 특징적인 표현이 없기 때문에 공간적 배경의 인식이 더욱 어렵다(Jang et al., 2005; Lee et al., 2000). 따라서 본 논문에서는 공간적 배경과 함께 사용될 수 있는 조사와 조사를 활용한 패턴을 이용하여 공간적 배경을 인식하는 기법을 제안한다. 조사 패턴을 이용한 공간적 배경 인식 기법의 흐름은 <Figure 3>와 같다.



<Figure 3> Spatial Background Candidates Recognition using Pattern that Utilizing Postposition

먼저 전처리 된 소설 텍스트가 입력되면, 모든 문장을 어절 단위로 분리한다. 분리된 각 어절에 사용된 조사가 공간적 배경과 함께 사용할 수 있는 조사인지 확인한다. 공간적 배경과 함께 사용할 수 있는 조사에는 ‘에’, ‘에서’, ‘에의’가 있으며, 이는 국립국어원 표준국어대사전에 ‘앞말이 처소의 부사어임을 나타내는 격 조사’로 설명되어 있는 조사이다. 본 논문에서는 실험을 통해 ‘에의’의 경우 소설 텍스트 전체에서 사용되지 않음을 확인하였다. 실험의 결과 중 일부는 <Table 3>과 같다.

<Table 3> Examples of Postposition Use Frequency

	‘에’	‘에서’	‘에의’
A Warrior of Goguryeo 1 (고구려 무사 1)	693	231	0
Independence Notes of Korea 1 (대한독립기 1)	1149	363	0
Cleopatra (클레오 파트라 1)	1008	356	0

따라서 본 연구에서는 소설 텍스트 내에서 각 어절에 ‘에’와 ‘에서’가 모두 사용된 단어만을 공간적 배경 후보로 인식한다.

- (a) 그는 조선에서 온 사람이었다.
- (b) 나는 말을 타고 조선에 갔다.
- (c) 그는 자신의 주머니에 칼을 넣었다.
- (d) 그의 주머니에서 돈이 떨어졌다.

예를 들어, 소설 텍스트에 (a)와 (b)가 있을 경우, ‘조선’을 공간적 배경 후보로 인식한다. 또한 소설 텍스트에 (c)와 (d)가 있을 경우, ‘주머니’를 공간적 배경 후보로 인식한다.

그러나 한국어의 특성상 조사 ‘에’와 ‘에서’ 만을 사용하여 공간적 배경 후보를 인식할 경우, ‘주머니’와 같이 공간적 배경이 아니지만 공간적 배경 후보로 인식되는 단어가 있어 정확도가 떨어지는 문제점이 있다. 따라서 ‘주머니’와 같이 공간적 배

경이 아니지만 조사 ‘에’와 ‘에서’를 둘 다 사용하는 단어를 제거하기 위해 조사를 활용한 패턴을 적용한다. 먼저 소설 텍스트에서 공간적 배경 후보를 포함하며 ‘에’와 ‘에서’를 사용한 문장을 모두 추출한 뒤, 형태소 분석기를 이용하여 형태소 분석을 한다. 본 연구에서는 꼬꼬마 형태소 분석기를 이용하여 형태소 분석을 한다. 형태소 분석 결과 ‘그런’ 또는 ‘다시’와 같은 관형사나 부사로 태깅된 어절은 제거한다. 그 후 형태소 분석 결과를 확인하여 조사패턴에 맞지 않는 공간적 배경 후보를 제거한다. 본 연구에서 사용한 조사패턴은 공간적 배경 후보 뒤에 조사 ‘에’가 사용되며 바로 다음 어절이 ‘동사+어미’의 형태로 되어있는 패턴이다. 조사 패턴에 맞게 사용된 공간적 배경 후보의 경우, 소설 텍스트에서 얼마나 많이 조사 패턴에 맞게 사용되었는지 빈도수를 추출한다.

예를 들어, (b)는 ‘조선’이라는 공간적 배경 후보 뒤에 조사 ‘에’가 사용되었으며, 바로 다음 어절이 ‘동사+어미’의 형태로 이루어져 있으므로 공간적 배경 후보로 인식한다. 다른 예로, (c)는 ‘주머니’라는 공간적 배경 후보 뒤에 조사 ‘에’가 사용되었으나, 바로 다음 어절이 ‘명사+조사’의 형태로 이루어져 있으므로 공간적 배경 후보에서 제거한다.

### 3.4 빈도수 기반 공간적 배경 인식

국가명 사전과 조사 패턴을 이용하여 인식된 공간적 배경 후보들의 사용 빈도수를 비교하여 순위를 정한다. 그 중 소설 텍스트에서의 사용 빈도수가 가장 높은 공간적 배경 후보를 최종적으로 소설 텍스트 내에 나타나는 대표적인 공간적 배경으로 인식한다.

## 4. 실험 및 결과

### 4.1 실험 절차

본 논문에서 제안한 공간적 배경 인식 기법의 정확도를 확인하기 위해 한국어 역사 소설 30권으로

실험을 진행하였다. 한국어 역사 소설은 한국 역사 소설 18권과 외국 역사 소설의 한국어 번역본 12권으로 이루어져 있다. 실험에는 txt형식의 소설 파일을 사용하였으며, 정확한 실험을 위해 각 소설 텍스트에 포함되어있는 제목, 서지정보, 목차, 저자의 말, 부록 등 본문이 아닌 내용은 수동으로 제거하였다. 실험에 사용된 도서는 <Table 4>와 같다.

<Figure 4>에서 <Figure 8>은 한국 역사 소설 중 ‘고구려를 위하여 1’을 실험한 예시이다. ‘고구려를 위하여 1’의 텍스트가 입력되면 전처리 과정을 거쳐 모든 문장을 같은 형식으로 일치 시킨다. ‘고구려를 위하여 1’은 한국 역사 소설이기 때문에 한국 역사 국가명 사전을 이용하여 공간적 배경 후보를 인식하고, 빈도수를 추출한다. 또한 조사 패턴을 사용하여 조사 ‘에’와 ‘에서’를 둘 다 사용하며, 조사 ‘에’ 바로 다음 어절이 ‘동사+어미’의 형태로 이루어져 있는 공간적 배경 후보를 인식 및 빈도수를 추출한다. 인식된 두 공간적 배경 후보의 빈도수를 비교하여 최종적으로 소설 ‘고구려를 위하여 1’의 공간적 배경으로 ‘고구려’를 인식한다.

<Table 4> Book List for Experiment

Book Name	
Korean Historical Novels	Warrior of Goguryeo 1 (고구려 무사 1) For Goguryeo 1 (고구려를 위하여 1)
	Independence Notes of Korea 1~7 (대한독립기 1~7) Chusa Kim Jung-Hee 1~6 (추사 김정희 1~6)
	Baekbeom's Daily Record (백범일지) Hong Kil-Dong (홍길동전)
	Lady Mishil (미실) -
Foreign Historical Novels	History of Three States 1~10 (삼국지 1~10권) History of Three States (the middle volume) (삼국지(중))
	Cleopatra (클레오 파트라) -

아아! 피라성

고구려 보장왕 27년 9월 5일.

인시 무렵이었다. 아직 단장을 깨기에는 이른 시각.

<Figure 4> Part of Text '고구려를 위하여 1'

아아  
피라성  
고구려 보장왕 27년 9월 5일  
인시 무렵이었다  
아직 단장을 깨기에는 이른 시각

<Figure 5> Part Among the Result of Preprocess

부여 6  
고구려 660  
백제 20  
신라 126  
발해 1

<Figure 6> Part Among the Result by Use of Korean Historical Nation Dictionary

곳곳 1  
피라성 10  
고구려 4  
아래 3  
장안 3

<Figure 7> Part Among the Result by Use of Postposition Pattern

고구려 660  
신라 126  
백제 20  
자리 15  
피라성 10

<Figure 8> Part Among the Result Compare

<Figure 9>부터 <Figure 11>은 외국 역사 소설 ‘삼국지(중)’을 실험한 예시이다. ‘삼국지(중)’은 외국 역사 소설이기 때문에 세계 국가명 사전을 이용하여 공간적 배경 후보를 인식하는 것 외의 절차는 한국 역사 소설인 ‘고구려를 위하여 1’과 동일하다. 두 공간적 배경 후보의 빈도수를 비교하여 소설 ‘삼국지(중)’의 공간적 배경은 ‘형주’로 인식된다.

말린 8  
인도 1

<Figure 9> Result by Use of World Nation Dictionary

신야 2  
강하 8  
행주 14

<Figure 10> Part Among the Result by Use of Post-position Pattern

행주 14  
신야 12  
자리 12  
강하 8

<Figure 11> Part Among the Result Compare

### 4.2 실험 결과

<Table 5>는 본 논문에서 제안한 기법을 이용하여 한국어 역사 소설을 대상으로 공간적 배경을 인식한 결과이다. 소설에서는 공간적 배경이 여러 개로 나타날 수 있다. 따라서 본 논문에서는 인식한 공간적 배경이 해당 도서의 공간적 배경이 될 수 있는지에 따라 정확도를 확인하였다.

공간적 배경 인식 결과 총 30권의 한국어 역사 소설 중 20권의 공간적 배경이 일치하여 66.67%의 정확도를 보였으며, 그 중 한국 역사 소설의 경우 83.33%, 외국 역사 소설의 경우 41.67%의 정확도를 보였다.

## 5. 결 론

본 논문에서는 한국어 역사 소설 텍스트에서 공간적 배경을 인식하기 위해 사전과 조사 패턴을 적용한 기법을 제안하였다. 또한 제안한 기법의 정확도 확인을 위해 한국어 역사 소설 30권을 대상으로 실험한 결과 66.67%의 정확도를 보였다. 인식된 공간적 배경은 기존에 제시된 이야기 공간 시각화 기법과 같이 소설의 공간적 배경을 시각화 하는데 활용될 수 있다(Park et al., 2014). 또한 인식된 공간적 배경은 도서정보를 기반으로 한 도서 추천 큐레이션 서비스에 활용될 수 있다(Kwon et al., 2015).

소설의 경우 공간적 배경이 여러 개 나타나거나, 국가명 외에도 지명이나 건물 등이 공간적 배경으로 나타나는 경우가 있을 수 있다. 따라서 향후에는 소설에 나타나는 다양한 범위의 공간적 배경 인식

<Table 5> Part Among the Result of Spatial Background Recognition

Book Name	Spatial Background	Matching Result
Warrior of Goguryeo 1 (고구려 무사 1)	Goguryeo (고구려)	T
For Goguryeo 1 (고구려를 위하여 1)	Goguryeo (고구려)	T
Independence Notes of Korea 1(대한독립기 1)	Joseon (조선)	T
Independence Notes of Korea 2(대한독립기 2)	Joseon (조선)	T
Independence Notes of Korea 3(대한독립기 3)	Joseon (조선)	T
Independence Notes of Korea 4(대한독립기 4)	Korean Empire (대한제국)	T
Independence Notes of Korea 5(대한독립기 5)	Korean Empire (대한제국)	T
Independence Notes of Korea 6(대한독립기 6)	Korean Empire (대한제국)	T
Independence Notes of Korea 7(대한독립기 7)	Korean Empire (대한제국)	T
Lady Mishil(미실)	Silla (신라)	T
Baekbeom's Daily Record (백범일지)	Joseon (조선)	T
Chusa Kim Jung-Hee (1) (추사 김정희 (1))	Joseon (조선)	T
Chusa Kim Jung-Hee (2) (추사 김정희 (2))	Joseon (조선)	T
Chusa Kim Jung-Hee (3) (추사 김정희 (3))	Joseon (조선)	T
Chusa Kim Jung-Hee (4) (추사 김정희 (4))	Goguryeo (고구려)	F
Chusa Kim Jung-Hee (5) (추사 김정희 (5))	Goryeo (고려)	F
Chusa Kim Jung-Hee (6) (추사 김정희 (6))	Goryeo (고려)	F
Hong Kil-Dong (홍길동전)	Joseon (조선)	T
History of Three States 1 (삼국지 1권)	China (중국)	T
History of Three States 2 (삼국지 2권)	Geography (자리)	F
History of Three States 3 (삼국지 3권)	China (중국)	T
History of Three States 4 (삼국지 4권)	Geography (자리)	F
History of Three States 5 (삼국지 5권)	Fifty Thousand (오만)	F
History of Three States 6 (삼국지 6권)	Fifty Thousand (오만)	F
History of Three States 7 (삼국지 7권)	Fifty Thousand (오만)	F
History of Three States 8 (삼국지 8권)	Nanyang (남양)	T
History of Three States 9 (삼국지 9권)	Fifty Thousand (오만)	F
History of Three States 10 (삼국지 10권)	Fifty Thousand (오만)	F
History of Three States (the middle volume) (삼국지(중))	Xingzhou (형주)	T
Cleopatra (클레오 파트라 1)	Egypt (이집트)	T
Accuracy		66.67(%)

기법에 대해 연구하여 도서 흐름에 따른 다양한 배경의 변화를 제시하고 공간적 배경 인식 정확도를 향상시킬 것이다. 또한 한국어 소설의 공간적 배경 뿐 아니라 시간적 배경을 인식하고, 소설 속 등장인물과 연계하여 등장인물이 나타나는 장면의 배경을 파악하여 소설의 분위기나 사건에 대한 이해를 도와줄 수 있도록 하는 연구를 진행할 계획이다.

## References

- Jaber, M.J. and S. Saad, "NER in English Translation of Hadith Documents using Classifiers Combination", *Journal of Theoretical and Applied Information Technology*, Vol. 84, No.3, 2016, 348-354.
- Jang, H.S., K.C. Jung, J.K. Lee, and K.H. Park, "Recognition of Korean Place Names on the Internet by Using the Rules of Dictionary Use", *Korean Society For Internet Information*, Vol.6, No.1, 2005, 397-400.  
(장혜숙, 정규철, 이진관, 박기홍, "사전을 이용한 규칙 기반의 세분화된 한국어 지명 인식", *한국인터넷정보학회*, 제6권, 제1호, 2005, 397-400.)
- Kim, J.H. and H.C. Kim, "Automatically Registering Schedules from SMS Messages on Handheld Devices", *Korean Journal of Cognitive Science*, Vol.22, No.1, 2011, 1-18.  
(김재훈, 김형철, "휴대전화에서 단문 메시지로부터 일정 자동 등록", *한국인지과학회*, 제22권, 제1호, 2011, 1-18.)
- Kwon, H.I., Y.B. Na, M.O. Yu, and K.S. Choi, "A Study of a Personalized Curation Service and Business Model based on Book Information", *Journal of Information Technology Services*, Vol.14, No.1, 2015, 251-262.  
(권혁인, 나운빈, 유미옥, 최광선, "도서정보 기반의 고객 맞춤형 큐레이션 서비스 및 비즈니스 모델 연구", *한국IT서비스학회지*, 제14권, 제1호, 2015, 251-262.)
- Lee, K.H., J.H. Lee, M.S. Choi, and G.C. Kim, "Study on Named Entity Recognition in Korean Text", *Proceedings of the 12<sup>th</sup> Annual Conference on Human and Cognitive Language Technology*, 2000, 292-299.  
(이경희, 이주호, 최명석, 김길창, "한국어 문서에서 개체명 인식에 관한 연구", *제12회 한글 및 한국어정보처리 학술대회*, 2000, 292-299.)
- Park, S.R., K.S. Lee, S.W. Ryu, and H.B. Kim, "Story Space Visualization Using Natural Language Processing", *The HCI Society of Korea*, 2014, 62-63.  
(박수란, 이기석, 류성원, 김현빈, "자연어 처리를 이용한 이야기 공간 시각화", *한국HCI학회 학술대회*, 2014, 62-63.)
- Prasad, G. and K.K. Fousiya, "Named Entity Recognition Approaches : A Study Applied to English and Hindi Language", *International Conference on Circuit, Power and Computing Technologies(ICCPCT)*, 2015, 1-4.
- Srinivasagan, K.G., S. Suganthi, and N. Jeyashenbagavalli, "An Automated System for Tamil Named Entity Recognition Using Hybrid Approach", *International Conference on Intelligent Computing Applications (ICICA)*, 2014, 435-439.
- Todorovic, B.T., S.R. Rancic, I.M. Markovic, E.H. Mulalic, and V.M. Ilic, "Named Entity Recognition and Classification using Context Hidden Markov Model", *9th Symposium on Neural Network Applications in Electrical Engineering*, 2008, 43-46.
- Zhang, Y.J. and T. Zhang, "Me-Based Chinese Person Name and Location Name Recognition Model", *International Conference on Machine Learning and Cybernetics*, Vol.6, 2007, 3442-3447.



## ◆ About the Authors ◆



**Seung-Hoon Kim (edina@dankook.ac.kr)**

Seung-Hoon Kim received his Ph.D. degree in Computer Science and Engineering from Pohang University of Science and Technology (POSTECH), Korea in 1998. Dr. Kim is currently a professor of Dept. of Applied Computer Engineering, Dankook University, Korea since 2001. From 1989 to 1990 he was a member of technical staff in Electronics and Telecommunications Research Institute(ETRI), Taejon, Korea. From 1991 to 1993 he was a member of technical staff in POSDATA, Seoul, Korea. His current research interests include data computing and networking, IoT, distributed systems, and etc.



**Seo-Hee Kim (rlatjgml38@naver.com)**

Seo-Hee Kim received the BE degree in Multimedia Engineering from Dankook University, Korea in 2015. She is presently a master student at the Dept. of Computer Science, Dankook University since 2015 under the supervision of Prof. Seung-Hoon Kim in Data Computing and Networking Lab. Her current research interests include IoT and data computing.