

해상교통관제 정보(음성) 수집 시스템 설계 및 구현

오정석, 권현수, 이효동, 김창묵, *김형훈
경북대학교 컴퓨터학부, 삼성전자

uh04049@naver.com, nowwater0225@gmail.com, good79797979@naver.com,
asde130@naver.com, pastelom@naver.com

Design and Implementation of Sea Traffic Control Information (Voice) Collection System

Jeong-Seok Oh, Hyeon-Su Gwon, Hyo-Dong Lee, Chang-Muk Kim,
Hyung-Hoon Kim

Dept of Computer Science, Kyungpook National University, Samsung
Electronics

요 약

현재 선박관제시스템에서는 관제정보들을 관제사가 직접 수기로 작성하여 관리하고 있다. 그 과정에서 관제사들은 번거로움을 느끼고 부정확한 정보 전달을 유발하기도 한다. 이에 본 논문은 관제사들이 구두로 녹음한 내용을 자동으로 문자로 저장하고, 저장된 선박 정보를 다시 음성 및 문자로 확인해볼 수 있는 시스템을 구현한다. 본 논문을 통해 관제사들이 보다 편리하고 정확하게 관제 정보를 관리할 수 있기를 기대한다.

1. 서론

해상교통관제 시스템에서는 선박의 안전과 원활한 통항을 목적으로 선박 자동 식별 장치(AIS), CCTV, 무선전화 등을 통해 선박에 관제 정보를 제공한다[1]. 현재 해상교통관제 시스템의 경우 관제사가 관제 정보를 직접 수기로 적어 관리하고 있다. 하지만 수기로 관제 정보를 관리하는 방식은 관제사의 불편을 유발할 뿐만 아니라 정보 관리의 정확성 또한 보장하기 어렵다.

선박의 항행 안전을 위해 교통, 보안 정보를 제공하는 해상교통관제 시스템에서 정보의 정확성은 매우 중요한 부분을 차지한다. 최근 정부에서는 해양관제 안전사고를 예방하고자 개발지원 및 정책적 방안들을 마련하고 있다. 해양경찰청은 선박 관제에 정확성을 높이고자 2020년 6월 4일부터 ‘선박 교통관제에 관한 법률’을 제정해 관제사, 선장, 선박 소유자의 책임과 의무를 강화하고 관련 기술 개발지원에 나서고 있다[2]. 이런 현황 속에서 수기로 관리되고 있는 많은 관제 정보들을 보다 안전하고 정확하게 관리할 방안을 마련해야 한다. 정책적으로 관제사에게 더욱 정확하게 정보를 관리할 책임을 강화할 뿐만 아니라 시스템상에서 기술적인 보완을 꾀할 필요가 있다.

이에 본 논문에서는 관제사가 음성녹음을 통해서 손쉽게 관제 정보를 관리할 수 있는 시스템을 설계하고 이를 구현하였다. 그리고 관제사는 본 시스템을 통해 저장된 선박 정보를 쉽게 조회해 볼 수 있다. 본 논문에서는 2장에서 관련 연구를, 3장에서는 설계, 4장에서는 구현 내용을 소개한다. 그리고 5장에서는 구현을 통한 최종 결론을 맺는다.

2. 관련 연구

2.1 STT(Speech-to-Text)

STT(Speech-To-Text)란 음성인식의 한 분야로서 사람의 음성언어를 컴퓨터가 해석하여 문자 데이터로 변환하는 처리를 의미한다[3]. 키보드와 같은 입력장치에 의한 입력이 아니라 사람의 음성을 이용하여 문자를 입력하는 방식으로 챗봇(Chat Bot), 인공지능 비서와 같은 다양한 기술개발에 활용되고 있다. STT는 사람의 음성을 HMM(Hidden Markov Model)과 같은 기계적인 알고리즘을 이용하여 문자로 변환하며, 최근에는 딥러닝(Deep Learning) 알고리즘을 사용하여 정확도가 더욱 향상되고 있다. STT에 이용되는 대표적인 딥러닝 알고리즘으로는 DNN(Deep Neural Network)이 있다.

2.2 TTS(Text-to-Speech)

TTS(Text-To-Speech)란 글자, 숫자, 기호와 같은 문자 데이터를 사람의 음성으로 변환하는 것을 말한다[4]. TTS는 문자를 읽는 데에 어려움이 있는 사람(유아, 시각 장애인 등)들이 음성을 통해 문자를 인식할 수 있게 해준다는 점에서 아주 효과적인 정보 제공기술이라고 할 수 있다. TTS 역시 기계적 알고리즘을 이용하여 문장을 분석하고 음성을 합성한다. 대표적인 알고리즘으로는 HMM(Hidden Markov Model)이 있다. 최근에는 DNN(Deep Neural Network) 알고리즘과 같은 딥러닝 알고리즘을 이용하여 적은 데이터만으로도 사람의 음성과 유사하고 자연스러운 합성음 제작이 가능해지고 있다.

2.3 형태소 분석

형태소란 언어에서 ‘최소 의미 단위’를 뜻한다. 최소 의미란 어휘적 의미, 문법적 의미를 모두 포함한다[5]. 형태소 분석이란 어절, 문장들을 최소 의미 단위인 형태소로 분할해내는 과정이다. 본 논문에서는 입력된 관제 정보 텍스트에서 형태소 분석 기술을 사용하여 관제 관련 단어를 추출한다. 형태소 분석 기술에는 모든 형태소를 하나의 벡터로 인식하는 딥러닝 기술이 포함된다[6]. 또한, 모든 단어와 구문, 논리적 표현은 하나의 벡터로 인식된다. 신경망은 두 개의 벡터를 하나의 벡터로 합치는 과정을 통해 정확하게 문장의 구조를 이해하고 해석할 수 있도록 도와준다. 그림 1은 글자 수가 적을 때는 다른 단어로 변화할 가능성이 높아 불확실성이 큰 편이지만, 점점 길어질수록 연속되는 문자들의 불확실성이 감소하여 특정 단어로 매핑될 확률이 커짐을 나타낸다.

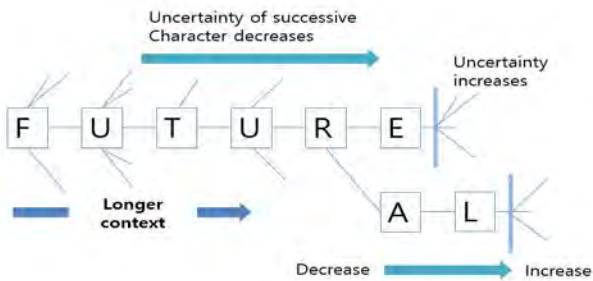


그림 1. 형태소 분석 과정

2.4 Blob(Binary large object)

Blob(Binary large object)은 데이터베이스 관리 시스템의 하나의 Entity로서 저장되는 이진 데이터 모임 형태이다[7]. 숫자나 문자는 정형화된 짧은 데이터인 것에 반해 그림, 사진과 같은 미디어 데이터는 비정형의 긴 데이터 형식이다. 멀티미디어와 같은 큰 데이터를 저장하기 위해서 사용하는 하나의 데이터 타입 형식을 Blob이라고 칭한다[8]. 본 논문에서는 사용자의 음성 오디오 파일 데이터를 Blob형식으로 변환하여 서버에서 클라이언트로 전송할 때 사용하였다.

3. 설계

3.1 시스템 구조도

그림 2는 본 논문의 웹 애플리케이션의 핵심적인 기능을 나타낸 시스템 구조도이다. 본 애플리케이션의 주요 기능은 크게 관제기록 조회와 관제 정보 녹음 기능, 이 두 가지이며 부가적으로 홈 화면 하단에는 해상교통관제 통계량을 연별, 월별, 일별로 확인해볼 수 있는 기능을 제공한다.

해상관제기록 조회 기능의 경우 서버 데이터베이스에 저장되어있는 해상관제기록들을 사용자가 선택한 조건에 부합하는 기록들만 선별하여 조회 할 수 있다.

관제 정보 녹음 기능은 STT(Speech-To-Text) API 기능을 사용해 사용자의 음성을 텍스트로 전환하여 해상관제 정보 형식에 맞게 추출 후 데이터베이스 임시저장 테

이블에 저장하는 역할을 한다. 추출과정이 모두 끝난 관제 정보 단어들은 관제사의 확인을 거치게 된다. 관제사에 의해 승인된 단어들은 최종적으로 데이터베이스에 저장되며 승인되지 않은 단어들은 데이터베이스에 저장되지 못하고 임시 테이블에서 삭제된다. 또 듣기 기능의 경우 TTS(Text-To-Speech) API 기능을 사용해 저장되어있는 텍스트를 음성으로 출력시켜준다. 데이터베이스의 테이블 구성은 다음 절에서 언급한다.



그림 2. 시스템 구조도

3.2 데이터베이스 설계

표 1은 본 논문의 ‘해상교통관제 정보 시스템’에 사용되는 데이터베이스 테이블의 구성이다. ‘ship’이라는 데이터베이스 안에 두 개의 테이블이 있으며, 테이블은 임시 정보를 저장하는 ‘temp’ 테이블과 관제사에 의해 승인된 실제 데이터를 저장하는 ‘record’ 테이블이 있다. 두 테이블 모두 동일한 구성을 가진다.

‘record’ 테이블에는 STT를 통해 얻은 데이터가 관제사가 ‘승인’ 또는 ‘삭제’를 선택하여 최종 결정을 짓기 전에 임시로 저장되는 곳이다. ‘record’ 테이블은 관제사가 임시 데이터를 ‘승인’할 경우 ‘temp’에 있던 데이터가 이동하여 저장되는 곳이다.

필드명	유형	필수/선택	설명
id	Int	필수	기본키
ship_name	Varchar	필수	선박명
weight_ton	Int	필수	총 톤수
ship_direction	Varchar	필수	입/출항
port_position	Varchar	필수	외/내항
date	Varchar	필수	날짜
time	Varchar	필수	시간

표 1. 데이터베이스 테이블

3.3 기능 흐름도

이 절에서는 시스템의 기능이 어떤 흐름으로 실행되는지 자세한 과정을 설명한다. 관제 정보조회기능은 해상관제기록 조회 페이지에서 조건을 선택해 조회 버튼을 클릭하게 되면 웹 서버는 데이터베이스의 ‘record’ 테이블 정보들을 요청한다. 데이터베이스는 웹 서버에게 전체 정보들을 제공을 해주고, 서버는 사용자가 선택한 조건들에 부합하는 정보들만을 추출하여 화면에 출력한다.

음성녹음기능은 관제사가 녹음하는 경우 웹 서버에서 실행되는 STT(Speech-To-Text) API가 입력된 음성 데

이터를 텍스트로 변환한다. 그 후 서버는 텍스트에서 주요 관제단어들만을 추출하고 추출 단어들을 데이터베이스 'temp' 테이블에 저장 요청한다. 올바른 데이터가 변환됐다고 판단이 되어 관제사가 승인할 시 웹 서버는 'temp' 테이블에 있는 데이터를 삭제 요청한 후 'record' 테이블에 데이터 저장을 요청한다. 올바른 단어가 아닐 시 관제사는 삭제 요청을 진행하고 이에 서버는 'temp' 테이블의 데이터를 삭제 요청하여 데이터를 완전히 지운다.

변환된 데이터가 성공적으로 'temp' 테이블에 저장됐는지 확인하기 위해 사용자는 듣기 기능을 사용할 수 있다. 듣기 요청 시, 웹 서버는 해당하는 데이터가 저장되어있는지 데이터베이스에 요청하고 데이터베이스는 이를 확인 후 Temp 테이블에 저장되어있는 데이터를 웹 서버에 제공한다. 데이터를 제공받은 서버는 TTS(Text-To-Speech) API 기능을 실행하여 텍스트데이터를 음성 파일로 변환 후 해당 음성 파일을 재생시켜 관제사가 정보를 음성으로 확인할 수 있게끔 기능을 제공한다.

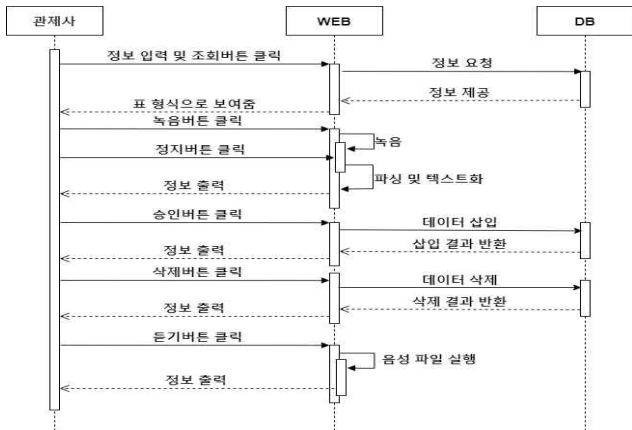


그림 3. 기능 흐름도

4. 구현

본 논문에서는 음성인식 기술을 활용하기 위해 구글 STT API를 활용했으며 시스템 서버 구현을 위해 Node.js Express를 사용하였다. STT의 과정은 크게 전처리, 패턴 인식, 후처리 총 3개로 나눌 수 있다. 그 중 핵심과 정인 패턴 인식 과정에는 다양한 종류의 알고리즘이 사용되는데 본 논문에서 사용한 API는 딥러닝 기반으로 DNN(Deep Neural Network) 알고리즘이 중점적으로 사용된다. DNN은 입력층과 출력층 사이에 그림 4에서 볼 수 있듯이 여러 개의 은닉층들로 이루어진 인공신경망이다. 일반적인 인공신경망처럼 복잡한 비선형 관계들을 모델링할 수 있다. 이 때, 추가계층들은 점진적으로 모여진 하위 계층들의 특징을 규합시킬 수 있기 때문에 다른 인공신경망에 비해 더 적은 수의 유닛(node)만으로도 더 복잡한 데이터를 모델링하는 것이 가능하다. DNN은 표준 오류 역전파 알고리즘을 통해 학습시킬 수 있으며 데이터의 양이 많아질수록 성능이 계속해서 향상될 수 있다는 측면에서 음성 인식 기술에 적합하다고 볼 수 있다.

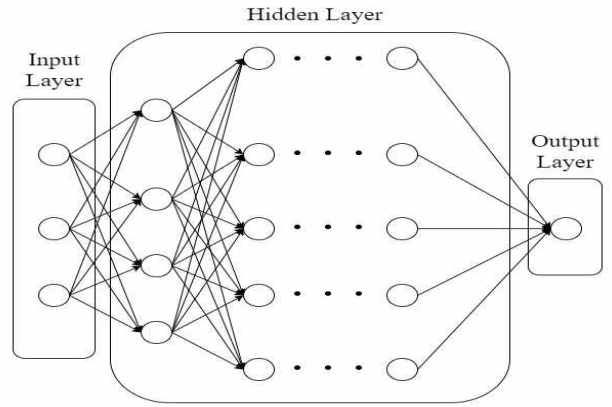


그림 4. DNN 구조

더불어 인식 정확도를 높이기 위해 구글 STT API의 음성 적응 기능을 추가 사용하였다. 음성 적응 기능이란 음성으로 감지된 문장에서 선택 관제에 자주 사용하는 단어를 음성 컨텍스트 힌트로 사용하여 추출 단어의 음성인식 정확성을 높이는 기술이다. 가령 '외항'이라는 단어를 녹음하는 경우 '왜항', '웨항'과 같은 잘못된 텍스트 변환을 막기 위해 주요 단어들을 음성 적응시켰다.

텍스트 변환 후에는 Soynlp 라이브러리를 통해 형태소 분석 과정을 거친다. Soynlp 라이브러리는 사용자 사전을 따로 구성해야하는 수작업을 생략할 수 있는 비지도 학습 형태소/품사 분석기이다. 단어 추출에는 'Cohesion Score'라는 확률 점수를 계산해 분석하게 되는 데, 이는 단어의 문맥을 살펴보고 다음 글자의 확률 상승/하강에 따라 단어의 경계를 판단하는 방식이다. '경북대학교는 대구에 위치해 있다.'라는 문장의 경우 '경'글자를 보고 '북'이 다음 글자일 것이라고 추측할 수 있는 확률은 매우 낮다. 하지만 '경북대'라는 글자를 보고 다음 글자가 '학'일 것이라고 추측하는 것은 비교적 확률이 높아진다. 그리고 '경북대학교'라는 단어 경계 이후 '는'을 예상하는 확률은 다시 작아지게 된다. 이렇듯 'Cohesion Score'를 통해 단어 경계를 판단하고 문장 전체에서 비지도 학습 방법으로 단어를 추출하게 된다.

$$Cohesion(c_{1:n}) = \sqrt[n-1]{\prod_{i=1}^{n-1} P(c_{1:i+1}|c_{1:i})}$$

수식 1. Cohesion Score 계산식

$$Cohesion('경북대학교') = p('경' \rightarrow '경북') * p('경북' \rightarrow '경북대') * p('경북대' \rightarrow '경북대학') * p('경북대학' \rightarrow '경북대학교')$$

수식 2. '경북대학교' 단어 Cohesion Score 분석

추가로 Soynlp 라이브러리의 경우 형태소 분석 시 단어 점수를 설정하여 분석할 수 있다. 단어점수는 다른 단어들보다 더 큰 점수를 부여할 시 해당 단어를 토큰라이저에서 하나의 단어로 더 쉽게 잘라낼 수 있다. 이 점을 활용하여 선택명, 총톤수, 입/출항, 외/내항, 날짜, 시간과 같은 주요 관제단어들에 단어점수를 설정한 후 Soynlp 라이브

러리 기능을 사용해 형태소를 분석한다. 그 후 완전하게 형태소로 분석되지 않은 단어들에 한해 2차 문자열 분석을 통해 요구하는 주요 선박 정보들을 모두 추출한다.

그림 5는 음성을 녹음해 데이터베이스에 정보를 저장하는 페이지이다. ‘record’ 버튼을 눌러 녹음을 시작하며 녹음을 끝낼 시 ‘stop’ 버튼을 눌러 종료한다. ‘stop’ 버튼을 누르게 되면 서버에서는 텍스트 변환, 형태소 분석, 단어 추출과정을 거쳐 웹페이지에 주요 관제 데이터를 출력한다. 관제사는 추출된 단어들을 확인해 알맞은 단어가 추출됐을 경우 ‘승인’ 버튼을 눌러 데이터베이스에 저장할 수 있다. 만약 잘못된 단어가 추출됐을 경우 삭제 버튼을 눌러 재녹음한다. 그리고 관제사는 듣기 버튼을 클릭하여 추출된 단어를 음성으로 확인해 볼 수도 있다.



그림 5. 음성녹음 페이지

그림 6는 조회 기능을 구현한 페이지이다. 조회는 선박명, 입/출항, 조회 시작 날짜, 조회 시작 시각, 조회 끝 날짜, 조회 끝 시간 기준으로 검색해볼 수 있으며 데이터베이스에 저장된 정보 중 검색 조건에 부합하는 정보들만을 선별해 웹페이지에 출력한다. 그림 7은 조회 결과를 출력한 웹페이지이다.



그림 6. 관제 정보 조회

3 Search listed

선박명	톤수	입항명	외사명	날짜	시간
백동호	1000	입항	내항	2020-07-11	20:00
백동호	192	출항	외항	2020-08-09	23:17
백동호	130	출항	외항	2020-08-09	23:17

그림 7. 관제 정보 조회 결과

5. 결론

본 논문에서는 선박 관제 정보를 관리하기 위해 딥러닝 알고리즘에 기반한 STT, TTS 기술을 이용하여 웹 애플리케이션을 구현하였다. 관제사가 직접 하던 작업을 음성 인식 기술을 적용해 자동화하여 데이터베이스에 관리함으로써 데이터의 정확성과 작업 효율성을 향상시킬 수 있을 것으로 예상된다. 또한, 웹 애플리케이션을 통해 실시간으로 관제 정보를 조회 및 추가할 수 있는 인터페이스를 제공한다. 따라서 관제사는 손쉽게 해당 인터페이스를 사용해 시스템을 관리할 수 있을 것으로 기대한다. 더불어 웹 페이지를 사용할 수 있는 환경이라면 추가 비용 없이 시스템을 손쉽게 보급할 수 있을 것이라 예상된다. 추후 한국어 인터페이스를 제공하고, 지역 사투리 인식률을 개선한다면 더 나은 서비스를 제공할 수 있을 것이다.

본 논문은 해양수산부 실무형 해상물류일자리지원사업의 지원을 통해 수행한 ICT멘토링 프로젝트의 결과물입니다.

참고문헌

- [1] 오승희, 김병두, 이병길, “해상교통관제 시스템과 항공교통관제 시스템의 비교분석”, 국가과학기술정보센터 공동 학술대회 논문집, pp. 201-203, 2013.
- [2] 일간경기, “해경, 6월 4일 선박교통관제 전문 법령 시행.. 안전사고 예방 기능 강화”, <https://www.1gan.co.kr/news/articleView.html?idxno=195276>(2020.0527)
- [3] 민소연, 이광형, 이동선, 류동엽, “한국어 특성 기반의 STT 엔진 정확도를 위한 정량적 평가방법 연구”, 한국산학기술학회논문지, 제21권, 7호, pp. 699-707, 2020.
- [4] 김태권, 김봉완, 최대림, 이용주, “안드로이드 OS 기반 한국어 TTS 서비스의 설계 및 구현”, 한국콘텐츠학회논문지, 제12권, 제1호, 2020.
- [5] Wikipedia - 형태소, <https://ko.wikipedia.org/wiki/형태소>
- [6] 김경선, <[인공 지능 기반 언어 처리 기술 - 자연어 대화 인터페이스를 중심으로]>, 《새국어생활》
- [7] WikiPedia - Blob, “https://ko.wikipedia.org/wiki/바이너리_라지_오브젝트”
- [8] 최기호, 강은지, 강현철 “멀티미디어 데이터의 효율적 저장을 위한 BLOB 타입의 다양화”, 정보과학논문지(B) 제22권 제10호, 1995,10