Smart tourism chatbot system using Multi-domain Tourism Information DST

Hoon-chul Kang
Dept. of Smart Electronic Control
Jeju Campus of Korea Polytechnic
Jeju, South Korea
kang9070@kopo.ac.kr

Ki-Beom Kang
National federation of fisheries cooperatives
Jeju, South Korea
kkb8671@gmail.com

Dong-Hyun Kim
Dept. of Telecomm. Eng.
Jeju National University
Jeju, South Korea
ehdgusdh@gmail.com

Myeong-Cheol Jwa
School of Computer Science and Eng
Korea University of Technology and Education
Cheonan-si, South Korea
jmc0407@koreatech.ac.kr

Tae-Seung Ko
Dept. of Telecomm. Eng.
Jeju National University
Jeju, South Korea
tsko7777@jejunu.ac.kr

Jeong-Woo Jwa
Dept. of Telecomm. Eng.
Jeju National University
Jeju, South Korea
lcr02@jejunu.ac.kr

***Abstract*— The smart tourism service provides tourists with travel planner services and tour guide services for easy and convenient travel throughout the entire travel process. In this paper, we develop the AI-based chatbot service using a pretrained language model (PLM) and provide tourism information so that tourists can make their travel plans. The proposed chatbot system consists of the DST server, the Neo4J graph DB and MySQL DB servers, and the natural language generation (NLG) server. The dialogue state tracking (DST) server understands the intention of tourists' questions to overcome the shortcomings of the previous rule-based chatbot system [7]. We define the domains and slots of the tourism information DST model with the 4W1H method and develop the dataset [12] for transfer learning of the SOM DST model [14]. The Neo4J and MySQL web servers search tourism information from the tourism information knowledgebase and the smart tourism information system, respectively. The NLG server provides the searched tourism information to the smart tourism app.**

Keywords— AI-based Chatbot, Smart Tourism, Dialogue State Tracking, Pre-trained Language Model, Tourism Information Knowledgebase.

I. Introduction

Smart tourism provides the travel planner and tour guide services to tourists using the Internet of Things (IoT), communication infrastructure, big data, artificial intelligence (AI), and AR/VR, which are the main technologies of the 4th industrial revolution. The travel planner service should be able to create a personalized itinerary including reservations for tourists. A smart tour guide service should be able to provide tourists with the same quality of service as traveling with a professional tour guide according to their itinerary. One of the ways to easily and conveniently provide smart tourism services to tourists is to use an artificial intelligence (AI)-based chatbot system [1][2]. In this paper, we develop the tourism information chatbot service that uses multi-domain tourism information dialogue state tracking (DST) algorithm [4][5] and tourist information knowledgebase of Neo4J graph DB. In Section II, we introduce Android and React smart tourism apps, the Text-to-Speech (TTS) server-based audio tour guide service, the rule-based chatbot system, and the tourism information DST dataset as previous studies to provide smart tourism services to tourists. In Section III, we describe the AI-based chatbot system using the multi-domain tourism information DST algorithm and the tourism information knowledgebase. Finally, in Section IV, we present our conclusions and future research.

II. Relate Works

A. Android and React Smart Tourism Apps

We have been developed the smart tourism Android and React apps that provide the personalized tour planner service and the tour guide service [6]. Figure 1(a) shows the main menu of the smart tour app. The smart tourism app consists of six menus: Best Tour Packages, My Travel Itinerary, Hallasan Beacon, Tourist Attractions, Accommodation, and Chatbot. Figure 1(b) shows the Best Tour Packages menu which provides travel products recommended by experts and local experts. Users can create their own travel itinerary in the My Travel Itinerary menu using the recommended travel products as shown in Fig. 1(c). In this paper, we develop the smart tourism chatbot system so that users can easily and conveniently select tourism information and use these to establish their travel plans. The Hallasan Beacon menu provides beacon-based tourism information and traveler safety services in Mt. Halla [7]. The Accommodation menu provides information such as restaurants, cafes, and lodging necessary for creating a travel plan. The smart tourism app also provides the tour guide service using walking, bus, and navi according to the their itinerary during travel with commercial map APIs. In addition, it provides real-time and weekly weather information for tourist destinations interworking with the weather API of the Korea Meteorological Administration.

(a) main menu (b) best tour packages (c) my travel itinerary

Fig. 1. Smart tourism app.

B. *Audio Tour Guide Service Based on the TTS Server*

We developed the Infodio™ app using the TTS server to provide location-based audio tour guide services to tourists at travel destinations according to the their itinerary [8]. The Infodio app consists of the tour guide service system, the wiki-based tour information collection system, the tourist information MySQL DB, the server-based TTS engine, and Android and IOS apps. Tourist location information is acquired through GPS coordinates, Beacon ID, WiFi ID, etc. The audio tour information system extracts the optimal tour information from the tour information DB and provides it to individual tourists. We provide natural language generation (NLG) service of smart tourism chatbot system using Infodio service.

C. *Rule-based Tourism Information Chatbot Service*

We developed the rule-based tourism information chatbot system [9] using the Khaiii (Kakao Hangul Analyzer III) morpheme analyzer [10] and the tourism knowledgebase of Neo4J graph DB [11]. To efficiently provide tourists with the tourism information provided by the smart tourism app as shown in Fig. 1, the tourism knowledgebase of the rule-based chatbot system was developed using the tourism information DB used in the app as shown in Fig. 2. The rule-based chatbot system provides an answer by understanding the intention of the question using the result of morpheme analysis using Khaiii for the user's question.

D. *Tourism Information Multi-domain DST datasets for Chatbot System*

The tourism information multi-domain DST dataset has 5 domains and 22 slots to provide appropriate tourism information by understanding the intention of the tourist's question [12]. The tourism information DST dataset was developed in the JSON format of the WOS dataset [13]. The tourism information DST dataset does not have DONTCARE and DELETE operations in value to improve the F1 score of the DST algorithm. The developed multi-domain tourism information DST dataset was evaluated using the selectively overwriting memory for dialogue state tracking (SOM DST) algorithm [14]. Experimental results show that the joint_goal_accuracy, turn_slot_accuracy, turn slot F1, operation accuracy are 0.9533, 0.9982, 0.9927, and 0.9992, respectively.

III. TOURISM INFORMATION CHATBOT SYSTEM USING THE MULTI-DOMAIN DST DATASET

A. *Smart Tourism Service Platform*

We develop the smart tourism service platform to provide tourist information, the recommended travel products, the personalized travel itinerary, and the tour guide services with smart tourism apps and chatbot services. Figure 2 shows the smart tourism service platform. The smart tourism service platform consists of the smart tourism information system for smart tourism services and the smart tourism chatbot system to provide chatbot services as shown in Fig. 2. The smart tourism information system has the tourist information DB, the recommended travel product DB, and the personalized travel itinerary DB built in MySQL DB. The Map view and the real-time weather information are provided through Naver, Daum, and Google Map APIs and the Korea Meteorological Administration weather API, respectively. The smart tourism service and chatbot service are provided to users through Android and React apps as shown in Fig. 1. The TTS server-based audio tour guide service is provided through the Infodio app according to the tourist's itinerary.

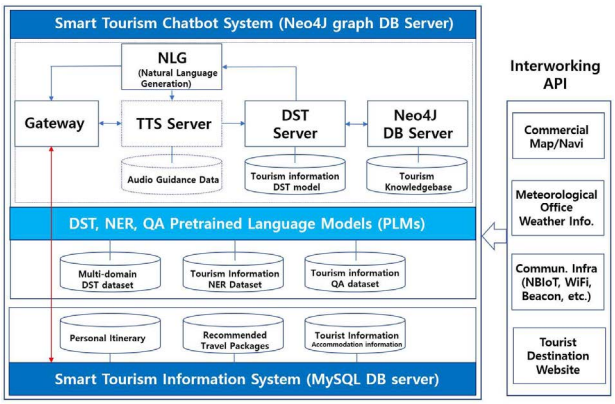


Fig. 2. The smart tourism service platform.

The tourism information chatbot system is developed by replacing the khaiii morpheme analyzer and the rule-based servers to the DST and policy servers. In the proposed chatbot system, the function to correct typos in user questions is the same as that of the rule-based chatbot system. In addition, the named-entity recognition (NER) [15] and the question answering (QA) datasets are developed for the smart tourism chatbot system using pre-trained language models. The tourism information DST model in the DST server is trained with the tourism information multi-domain DST dataset.

B. *Tourism Information Knowledgebase using Neo4J Graph Database*

The smart tourism chatbot system should be able to provide easy and convenient travel services by integrating with the smart tourism apps that provide tourist information, the recommended travel products, the personalized itinerary, and the tour guide service to individual tourists. In this paper, we replace the khaiii morpheme analyzer and rule-based chatbot system used to understand the intention of the user's question with the tourism information multi-domain DST model. The tourism information knowledgebase of the Neo4J graph DB should be newly designed and built according to the domain, slot, and value of the DST model. We create the tourism information knowledgebase so that nodes and relationships of the Neo4J graph DB can be set according to the domain and slot of the DST model. In the design and development of the new tourism information knowledgebase, the following items are considered based on the domain, slot, value of the DST model, the tourism information MySQL DB of the smart tourism information system, and the previous tourism information knowledgebase of Neo4J DB.

(1) As OSMU (One-source Multi-use), the tourism information chatbot service have to use the same contents of the smart tourism information system provided in the smart tourism app as shown in Fig. 1. The tourism information chatbot system provides answers to user questions with the tourism information knowledge base of the Neo4J graph DB, and provides detailed tourism information requested by users through MySQL DB.

(2) The tourism information chatbot service also have to use the same tourist destination classification system as that of the smart tourism app. Through this, the chatbot system can provide users with the same tourist information provided by the Tourist Attractions menu.
(3) The tourism information chatbot service should be linked with Instagram and YouTube services.
(4) Redesign domain, slot, and value of the DST model and additionally develop the related datasets to accurately understand the intention of tourist's questions.

The tourism information knowledgebase used in the rule-based chatbot system has 3 nodes of Area, Contents and Services. The Area node represents the location of a tourist destination in terms of administrative districts and districts, and is created based on the address of the tourist destination. The Area node is used to provide tourist information within the region requested by the user. The Contents node represents tourism information provided to users and is created with some tourism information data of MySQL DB in the smart tourism information system. Detailed tourism information requested by the user is provided in the smart tourism information system using the tourism information id provided by the Contents node. The Services node is composed of attribute data that can provide appropriate tourism information according to the intention of the user's question identified by the DST server. The attribute data of the Services node is made as shown in Table 1 according to the 4W1H method. The quality of the smart tourism chatbot system is determined by the performance of the DST model and the quality of the attribute data of the Contents node. We modify and redesign the attribute data of the Services node to provide personalized tourism information to tourists. Table 1 shows the modified slots of the tourism purpose domain using the 4W1H method.

TABLE I. THE MODIFIED ATTRIBUTE DATA OF THE SERVICES NODE

Who	Travel commentator, Alone, Friend, Family, Child, Couple, Colleague, Pet
What	Cultural tourism, Dark tourism, Sea experience, Wellness experience, etc.
When	Season(Spring, Summer, Autumn, Winter), Time(weekdays, weekends, now, morning, noon, evening, day, night), Duration(Festival, Flower, open time), etc.
Where	Category(Large, Medium, Small), Photo(Flowers, Sunrise/Sunset), etc.
How	Climbing, Walking, Driving, Biking, Personal transportation, etc.

C. *Neo4J Query Statement and Search Results from the Tourism Information Knowledgebase*

The DST server analyzes the user's question and determines the domain, slot, and value. Figure 3 shows the process of providing answers to user questions in the smart tourism chatbot system. The DST server generates domain, slot, and value for user questions by using the DST model that performed transfer learning on the SOM DST model with the tourism information multi-domain DST dataset. The DST server creates Neo4J query statements to search tourism information in the tourism information knowledge base according to domain, slot, and value of the DST model. Figure 4 shows an example of a Neo4J query statement according to the domain, slot, and value analyzed by the user's question in the DST server to search from the tourism information knowledgebase. The Neo4J query statement consists of 4 nodes of Area, Contents, Category, and Services and 3 relationships of IN_AREA, IN_CATEGORY and IN_SVC. The Category node provides not only tourism information but also lodging, restaurants, and travel products as chatbot services. The Neo4J graph DB server searches for tourism information according to the WHERE syntax of the Neo4J query statement and transmits the search result to the DST server according to the RETURN syntax. Regarding the Neo4J query example in Figure 4, the Neo4J DB server searches for tourist information matching Area node (= East of Seogwipo-si) and Services node (family, medium-class category=Oreum). The Neo4J DB server transmits 10 tourist information to the DST server in the order of likes (vcount) among the searched tourism information according to the RETURN syntax (tourist information id, tourist destination name, address, etc.). The DST server creates answer sentences to user questions according to the domain, slot, value of the DST model and the tourism information search result from the Neo4J graph DB server.

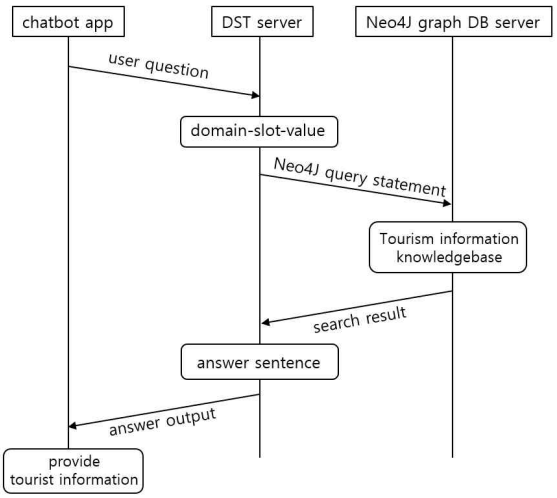


Fig. 3. the process of providing answers to user questions in the smart tourism chatbot system.

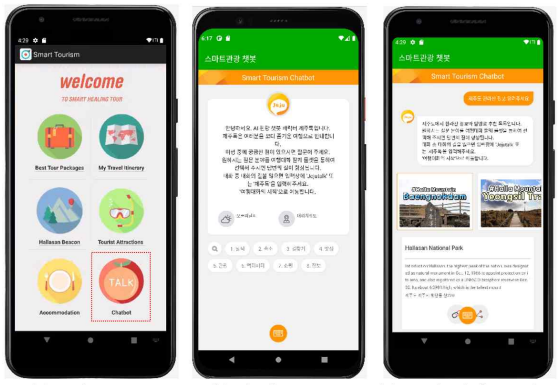
User : Please tell me which oreum is good to go with my family in the east of Seogwipo City.

DST analysis : ['Tourism information-Classification-Oreum', 'Tourist area-Area4-East of Seogwipo-si', 'Tourism purpose-Who-Family']

Neo4J query statement :

```
MATCH (a:Area)<-[:IN_AREA]-(c:Contents)
  - [:IN_CATEGORY]->(ca:Category{name:'4000000'}), (s:Service)<-[ :IN_SVC]-(c)
WHERE (a.area_div =~ 'East of Seogwipo-si')
  AND (s.where_medium =~ 'oreum') AND (s.who_family =~ 'family')
RETURN DISTINCT c.id AS id, c.name AS name, c.addr AS addr, c.tel AS tel,
  c.homepage AS homepage, c.latitude AS lat, c.longitude AS lot, c.desc AS desc,
  c.vcount ORDER BY c.vcount DESC LIMIT 10
```

Fig. 4. An example of the Neo4J query statement in the DST server.



(a) main menu (b) chatbot menu (c) tourist information

Fig. 5. The tourism information app screen in the smart tourism app.

Figure 5 shows the tourism information app screen in the smart tourism app. Figure 5(b) shows the main screen of the smart tourism chatbot service. Figure 5(c) shows that the chatbot system provides the tourism information of Mt. Halla retrieved from the tourism information knowledgebase as an answer to the user's question of "Please tell me information about Mt. Halla in Jeju Island?". The user can request detail tourism information to select tourist destinations to create his or her own travel plans. Figure 6 shows the process of providing the detail tourism information as that provided by the Tourist Attractions menu in the smart tourism information system. The smart tourism chatbot app requests medium-class category tourism information to the smart tourism information system using the tourism information id (c.cnt_id) that is composed of large-class, medium-class, and small-class categories.

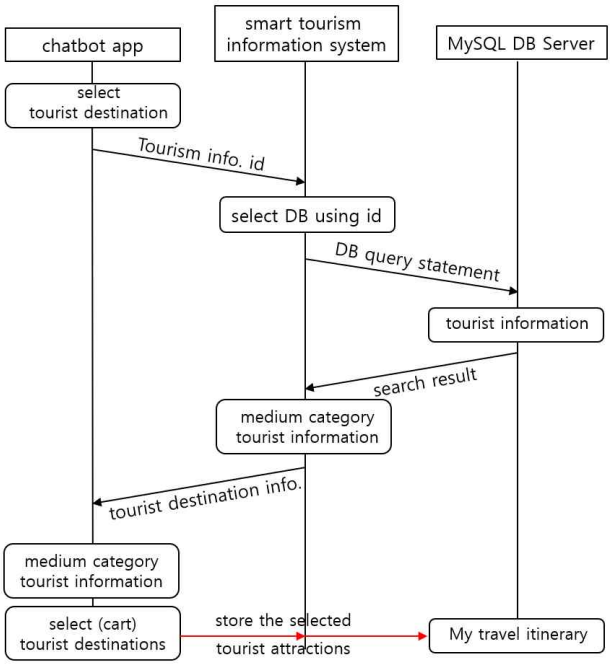
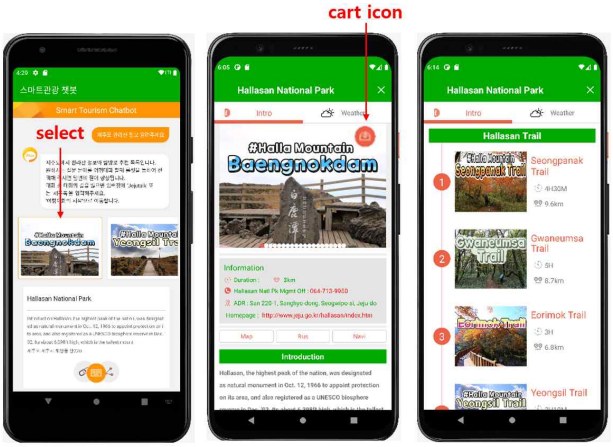


Fig. 6. The process of providing the detailed tourism information in the smart tourism information system.



(a) search result (b) tourist information (c) medium category

Fig. 7. The tourist information from the smart tourism information system.

Figure 7 shows the tourism information provided by the smart tourism information system in the chatbot app. When the user requests tourism information on Baengnokdam as shown in Fig. 7(a), the smart tourism information system retrieves tourism information from MySQL DB and provides it as shown in Fig. 7(b). The user can select tourism destinations with the cart icon in Fig. 7(b) to create his or her own travel plans and can check the selected tourist destinations in the My Travel Itinerary menu. The tourism information chatbot app provides tourism information on the 7 climbing courses of Mt. Halla as medium-class category tourism information along with Baengnokdam as shown in Figure 7(c).

IV. CONCLUSIONS AND FURTHER WORKS

Smart tourism services should be able to provide tourists with context-aware personalized tour planner services and tour guide services such as those with professional guides. Although there are still some inaccurate contents in the answers provided by the AI-based tourism information chatbot service to tourists, it will be able to provide smart tourism services in the future. In this paper, we propose the AI-based tourism information chatbot system that combined with the smart tourism apps to easily and conveniently provide smart tourism services to individual tourists. The proposed chatbot system uses tourism information multi-domain DST model and Neo4J graph DB to retrieve tourism information for context-aware personalized travel planner service. In addition, the chatbot system provides tourist destination information from the smart tourism information system so that users can select tourist destinations for the travel planner service.

The proposed tourism information chatbot system consists of gateway, the DST server, the Neo4J graph DB server, and the management server. The DST server uses the SOM DST model [14] to understand user's questions with the newly defined tourism information domain, slot, value. In this paper, we performed transfer learning by complementing the existing tourism information multi-domain DST dataset [12] according to the newly defined domain, slot, and value based on the 4W1H method. The supplemented DST dataset also does not use DONTCARE and DELETE operations to improve the performance of the DST model. The tourism information knowledgebase consists of the Area, Contents, and Service nodes and three relationships. Contents node is created using the tourism information DB of the smart tourism apps so the tourists can use smart tourism apps, Instagram, and YouTube services through the smart tourism chatbot service. In the future, We will develop the smart tourism chatbot service that provides the recommended travel packages, the personalized travel planner service, and the smart tour guide services. We also plan to improve the quality of smart tourism services by applying NER and QA models to the smart tourism chatbot system.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2021R1A2C1093283).

REFERENCES

[1] Davide.Calvaresi, et.al. ,"The evolution of chatbots in tourism: A systematic literature review" Information and Communication Technologies in Tourism, pp. 3-16, Jan. 2021.

[2] MARK ANTHONY CAMILLERI and CIRO TROISE, "chatbot recommender systems in tourism: a systematic review and a benefit-cost analysis", 8th International Conference on Machine Learning Technologies, pp. 1-10, March, 2023.

[3] Xipeng Qiu, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, Xuanjing Huang, “Pre-trained Models for Natural Language Processing: A Survey”, Science China Technological Sciences 63(10), pp.1872-1897, 2020.

[4] Hongshen Chen, Xiaorui Liu, Dawei Yin, Jiliang Tang,"A Survey on Dialogue Systems: Recent Advances and New Frontiers", ACM SIGKDD Explorations Newsletter, Vol. 19 Issue 2, pp. 25-35, December 2017.

[5] Léo Jacqmin, Lina M. Rojas-Barahona, Benoit Favre, ""Do you follow me?": A Survey of Recent Approaches in Dialogue State Tracking", Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp. 336–350, Sept. 2022.

[6] Hyun-Ji Cho, Jin-Yi Lee, Tae-Rang Park, Jeong Woo Jwa, "React Native and Android Mobile Apps for Smart Tourism Information Service to FITs", The International Journal of Internet, Broadcasting and Communication, vol.14, no.2, pp.63-69, 2022.

[7] Ko, Tae-Seung, Kim, Byeong-Joo, Jwa, Jeong-Woo, "Smart Tourism Information System and IoT Data Collection Devices for Location-based Tourism and Tourist Safety Services", International Journal of Advanced Culture Technology, Vol. 10 No. 1, pp.310-316, 2022.

[8] KiBeom Kang, JeongWoo Jwa, SangDon Earl Park, “Smart Audio Tour Guide System using TTS”, International Journal of Applied Engineering Research, pp.9846-9852, 2017.

[9] Dong-Hyun Kim, Hyeon-Su Im, Jong-Heon Hyeon, Jeong-Woo Jwa, “Development of the Rule-based Smart Tourism Chatbot using Neo4J graph database”, International Journal of Internet, Broadcasting and Communication, Vol.13, No.2, pp 179-186, 2021.

[10] Kakao khaiii(Kakao Hangul Analyzer III), https://tech.kakao.com/2018/12/13/khaiii/

[11] Neo4j graph database, https://neo4j.com/

[12] Myeong-Cheol Jwa, Tae-Seung Ko, Byeong-Joo Kim, Jeong-Woo Jwa, "Tourism Information Multi-domain Dialogue State Tracking Datasets for Smart Tourism Chatbot", International Journal of Intelligent Systemsand Applications in Engineering, vol.10, no.1S(2022), pp. 192-196, 2022.

[13] Park, Sungjoon, Jihyung Moon, Sungdong Kim, Won Ik Cho, Jiyoon Han, Jangwon Park, Chisung Song et al., “KLUE: Korean Language Understanding Evaluation.”, arXiv, 2021.

[14] Kim, Sungdong, et al. "Efficient dialogue state tracking by selectively overwriting memory." Proceedings of the 58th Annual Meetings of the Association for Computational Linguistics, 2020.

[15] Myeong-Cheol Jwa, Jeong Woo Jwa, “Development of Tourism Information Named Entity Recognition Datasets for the Fine-tune KoBERT-CRF Model”, International Journal of Internet, Broadcasting and Communication, Vol.14, No.2, pp 55-62, 2022.