Implementation of Danger Degree Calculation System for Public Safety Services

Hyunho Park
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
hyunhopark@etri.re.kr

Eunjung Kwon
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
ejkwon@etri.re.kr

Sungwon Byon
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
swbyon@etri.re.kr

Minjung Lee
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
minjunglee@etri.re.kr

Young Soo Park
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
yspark@etri.re.kr

Eui-Suk Jung
Police Science & Public Safety ICT Research Center
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
esjung@etri.re.kr

***Abstract*—Danger degree calculations are important to provide public safety services because crime response methods vary depending on danger degrees of crime scenes. This paper proposes a danger degree calculation system (DDCS) to calculate the danger degrees by using natural language processing (NLP) of emergency report data that summarize emergency calls. This paper also explains an implementation of DDCS by building a deep learning NLP model with KoBERT and visualizing a process of danger degree calculations. Danger degrees from the DDCS will help a constabulary to determine appropriate physical forces (e.g., devices and crime control actions) for police officers' responding to the crime scenes.**

Keywords*—*danger degrees, DDCS, KoBERT

I. INTRODUCTION

It is difficult for police officers to respond to crime scenes with appropriate physical forces [1-3]. The physical force can consist of manpower allocations (e.g., the number of police officers), device types (e.g., police sticks and taser guns), and crime control actions (e.g., arrest and shooting a taser gun) of the police officers that responds to crime scenes. Statistics of Korean National Police Agency (KNPA) [1] show that 1,395 police officers were wounded in South Korea of 2019 because the officers exert physical forces inappropriately. Study [2] explains that an appropriate manpower allocation of police officers is a difficult issue as crimes become more diverse and complicated. Study [3] discussed appropriate crime control actions by reviewing the precedents of inappropriate crime control actions of police officers with inappropriate devices.

Physical forces and system technologies to respond crime scenes have been studied [4-7]. KNPA establishes rule No. 550 to determine physical force grades (PFGs) that indicate device types and crime control actions to respond a crime scene [4]. The rule No. 550 defines criminals' activity grades (CAGs) from "adaptation" to "fatal attack" depending on severity of criminals' aggressive behaviors in a crime scene. The rule No. 550 also matches CAGs to PFGs that ranges from "cooperative control" to "high-risk physical force." Los Angeles Police Department developed predictive policing (PredPol) system for manpower allocation of police officers to prevent crimes [5]. The PredPol analyzes crime data that include types, time, and locations of crimes and predicts future crime locations based on the crime data. The PredPol allocates police officers to patrol the future crime locations. We proposed an architecture of emergency dispatch support system (EDSS) to respond occurred crime scenes [6]. The EDSS can analyze status of crime scenes based on emergency calls and crime data. The EDSS can generate response information for responding to the crime scenes. We also proposed a crime response system based on danger degree calculations (CRSDDC) system for helping the EDSS to analyze the status of crime scenes and to generate response information [7]. The CRSDDC system calculates danger degrees of crime scenes based on the emergency calls and crime data. The CRSDDC system generates physical force information depending on the danger degrees. Study [7] also proposed equations to calculate danger degrees by using the emergency calls and crime data.

In this paper, we propose a danger degree calculation system (DDCS) that can calculate danger degrees based on emergency report data that summarize emergency calls. In addition, we explain implementation of the DDCS. An architecture and equation to calculate danger degrees are developed by referring to [7]. The DDCS generates danger degrees of the DDCS by learning emergency report data based on KoBERT [8] that is a famous Korean language processing model. The DDCS visualizes performances of learning results and also visualizes a process of danger degree calculations.

The remainder of this paper is organized as follows. Section II illustrates an architecture of the DDCS. Section III explains an implementation of the DDCS. Section IV provides concluding remarks and future works of this paper.

II. DANGER DEGREE CALCULATION SYSTEM (DDCS) BASED ON EMERGENCY REPORT DATA

This section explains the DDCS and its implementation with emergency report data (ERD). Figure 1 shows architecture of the DDCS. The DDCS receives ERD and CAG labeling data from CAG labeler. The ERD are the data that recipients of emergency response centers summarize emergency calls into. The ERD include times, locations, crime types, and summarized information about crime scenes. The CAG labeler matches CAG labels (i.e., "adaptation," "passive resistance," "active resistance," "violent attack," and "fatal attack") to the ERD. The DDCS trains the ERD with CAG labels and calculates danger degrees. The DDCS visualizes for the train results, test results, and danger degree calculations.

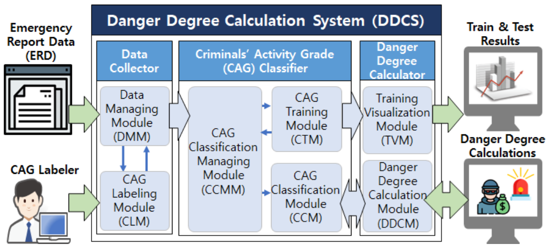


Fig. 1. Architecture of DDCS.

The CAGs in rule No. 550 [4] include "adaptation," "passive resistance," "active resistance," "violent attack," and "fatal attack." The "adaptation" means that a criminal does not show any aggressive behaviors. The "passive resistance" indicates that the criminal shows mild aggressive behaviors but does not threat people. The "active resistance" implies that the criminal may threat people but does not inflict physical damage to people. The "violent attack" means the state that the criminal can inflict physical damage to people. The "fatal attack" indicates that the criminal can inflict serious physical damage to people or can kill people.

The DDCS consists of a data collector, criminals' activity grade (CAG) classifier, and danger degree calculator. Data collector collects and manages ERD with a data managing module (DMM). A CAG labeling module (CLM) receives ERD from the data managing module and also receives CAG labels from the CAG labeler. The CLM matches CAG labels to the to the ERD by using inputs from the CAG labeler. The DMM sends the ERD to the CAG classifier.

The CAG classifier trains the ERD and classifies CAGs of the ERD. A CAG classification managing module (CCMM) receives the ERD from the data collector. The CCMM separates the ERD into the datasets for a train phase, validation phase, and test phase. The CCMM manages to train the ERD to generate CAG classification models. The CCMM gives a CAG training module (CTM) the data for a train phase, validation phase, and test phase. The CTM can learn the data with deep learning technologies. The CTM also generates the CAG classification models that can classify CAGs by using the ERD. The CTM gives the CAG the train and test results to the danger degree calculator. The CTM also gives the CAG classification models to the CCMM. The CCMM gives the CAG classification models to a CAG classification module (CCM), and the CCM classifies CAGs by receiving ERD from danger degree calculator. The CCM gives the classified CAGs to the danger degree calculator.

The danger degree calculator calculates danger degrees and visualizes the CAG train results, test results, and danger degree calculations. A training visualization module (TVM) of the danger degree calculator visualizes the CAG train results. A DDCS user can input ERD to a danger degree calculation module (DDCM). The DDCM gives the ERD to CCM. The CCM classifies a CAG for the information and then provides the classified CAG to the DDCM. The DDCM calculates danger degrees with the CAGs from the CAG classifier.

A danger degree (DD) calculated in the DDCM is defined as follows:

$$DD = \sum_{i=1}^{NC}\{CAS_i \times WA_i \times WW_i\}, \quad (1)$$

where NC, CAS_i, WA_i, WW_i are the number of criminals, activity score, age weight, and weapon weight of the *i-th* criminal, respectively. The CAS_i is determined depending on CAG_i (that is, the CAG of the *i-th* criminal). The CAS_i, is 1 when the CAG_i is "adaptation," 2 when the CAG_i is "passive resistance," 3 when the CAG_i is "active resistance," 4 when the CAG_i is "violent attack," and 5 when the CAG_i is "fatal attack." The WA_i and WW_i vary respectively depending on a criminal's age and weapon (e.g., a knife or gun). Criteria for WA_i and WW_i are determined by implementation methods. The next section describes implementation of the DDCS including the criteria for WA_i and WW_i.

III. Implementation of the DDCS

We implemented the DDCS based on the architecture of section II. The ERD in this paper were generated by our partner company that has experiences to use emergency call data. The partner company deidentified private information in the ERD. The ERD have information on time, location, and summary of 10,000 crime cases. DMM of data collector collects the ERD and sends the ERD to the CLM. We as a CAG labeler tags CAG labels (i.e., "adaptation," "passive resistance," "active resistance," "violent attack," and "fatal attack") to 10,000 crime cases in the ERD through CLM. CMM gives CAG classifier the ERD with CAG labels.

The CAG classifier generates CAG classification models by learning the ERD with KoBERT [8]. DMM splits the ERD into datasets for train phase, validation phase, and test phase with a ratio of 60:20:20. The DMM sends the datasets to CTM. The CTM learns the datasets for the train phase with KoBERT. Batch size of the datasets for the train phase is 32. BERT SentencePiece tokenizer is used to tokenize sentences of the ERD. KoBERT model with 0.3 dropout rate is used to train the ERD for classifying CAGs. Number of the epochs for training datasets is 10. An optimizer and cost function are respectively AdamW with 0.00005 learning rate and cross-entropy function. Clip gradient norm is used to prevent gradient exploding problems with 1 max norm of gradients. A scheduler of CTM is cosine scheduler with 188 warmup steps.

Results of training and validation phases can be showed via TVM of danger degree calculator. Figure 2 shows training accuracy (*train_acc*) and validation accuracy (*valid_acc*) for CAG classifications over training epochs. After the second training epoch, *tain acc* is much higher than *valid acc*.

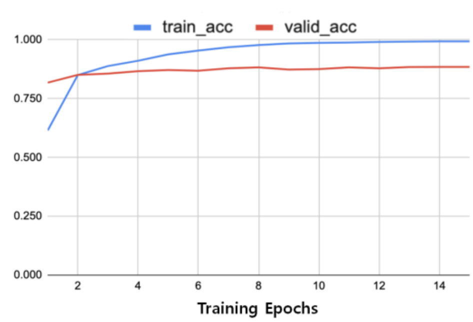


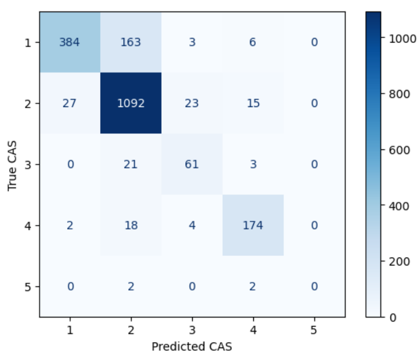
Fig. 2. Training accuracy and valid accuracy for CAG classifications over training epochs.

It means that train results are overfitted after the second training epoch. Thus, we use the CAG classification model of the second training epoch for test phase.

Figure 3 shows results of a test phase with the CAG classification model of the second training epoch. Figure 3 (a) illustrates a classification report that includes precision, recall, F1-score, and accuracy. In Fig. 3 (a), the numbers from 1 to 5 in the left column are *CAS* values that are matched to CAGs. Accuracy and weighted average F1-score of the test phase are respectively 0.86 and 0.85. Figure 3 (b) shows a complex matrix between true *CAS* values and predicted *CAS* values. When the *CAS* value is 5 (i.e., the CAG is "fatal attack"), F1-score is 0. Because the ERD that is matched to "fatal attack" are about 2% in all of the ERD. Therefore, the CTM cannot sufficiently learned cases of the CAGs that are "fatal attack."

	precision	recall	f1-score	support
1	0.93	0.69	0.79	556
2	0.84	0.94	0.89	1157
3	0.67	0.72	0.69	85
4	0.87	0.88	0.87	198
5	0.00	0.00	0.00	4
accuracy			0.86	2000
macro avg	0.66	0.65	0.65	2000
weighted avg	0.86	0.86	0.85	2000

(a) Classification report



(b) Complex matrix

Fig. 3. Results of a test phase for CAG classifications.

DDCM receives inputs of ERD from a DDCS user and calculates danger degrees for the ERD. The DDCM gives the ERD to CCM of the CAG classifier. The CCM classifies a CAGs from the ERD by using the CAG classification model of the second training epoch. The CCM sends the classified CAGs to the DDCM. The DDCM matches the CAGs to *CASs* as explained in section II. The DDCM extracts information on the *i-th* criminal' ages (CA_i), the *i-th* criminal' weapon (CW_i), and the number of criminals (*NC*). In this paper, we extract the CA_i, CW_i, and *NC* by simple keyword matching. The DDCM matches the CA_i to WA_i and also matches CW_i to WW_i.

Table I shows relationships among keywords, *CA*, *CW*, *WA*, *WW*, and *NC*. If there is a "청소년" (that is, "adolescent" in Korean) in ERD, *CA* is "teenager." If there is a "칼" (that is, "knife" in Korean) in ERD, *CW* is "short-range weapon." The number before "명" (that is, "the number of people) in ERD can be *NC*. If there is no keyword of Table I in ERD, *WA*, *WW*, and *NC* are set to 1.

TABLE I. RELATIONSHIP BETWEEN KEYWORDS, *CA*, *CW*, *WA*, *WW*, AND *NC*

Keywords	*CA* or *CW*	*WA*, *WW*, or *NC*
청소년, 중학생, 고등학생, 십대	*CA*= "teenager"	*WA* =1.2
칼, 단도, 과도	*CW*= "short-range weapon"	*WW*=1.1
각목, 파이프, 전기총, 가스총	*CW*= "mid-range weapon"	*WW*=1.2
총, 권총, 장총, 사냥총, 라이플, 화살	*CW*= "long-range weapon"	*WW*=1.3
Number 명	-	*NC*=*Number*

The DDCM can visualize danger degree calculations. Figure 4 explains a visualization example of a danger degree calculation. If a DDCS user inputs "청소년 2 명이 시비가 붙어 칼을 들고 서로를 협박한다" (that is, "Two adolescents threat each other with knives." in English), the DDCM calculates a *CAS* as 4 that is matched to a CAG as "violent attack". Based on the rule with Table I, *NC*, *WA*, and *WW* are derived as 2, 1.2, and 1.1, respectively. Therefore, the DD is calculated in DDCM is 10.56.

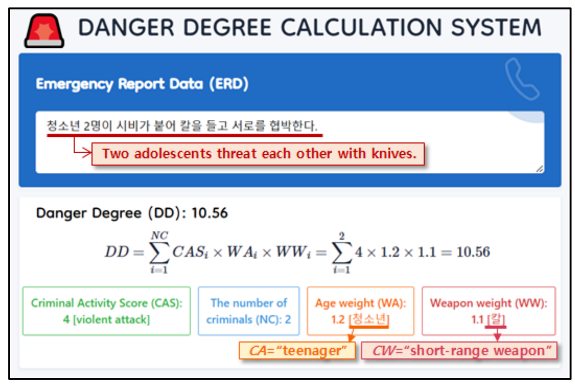


Fig. 4. Visualization example of a danger degree calculation.

IV. CONCLUSION

This paper proposes danger degree calculation system (DDCS) and explains implementation of the DDCS based on emergency report data (ERD). The DDCS learns ERD to classify criminals' activity grades (CAGs) with KoBERT and visualizes the learning results. The DDCS shows 0.86 accuracy to classify the CAGs. In addition, the DDCS receives ERD inputs form a DDCS user and visualizes calculations of danger degrees by using the ERD inputs.

Future works will improve the DDCS in perspective of implementation. As the first future work, we will propose criteria of physical forces based on the danger degrees. Purpose for the danger degree calculations is to give police officers appropriate criteria of the physical forces that are matched to the danger degrees. Thus, a method for physical force calculations needs to be proposed. As the second future work, we will provide extended rules to determine a criminal's age weight (*WA*) and weapon weight (*WW*). Table I shows a simple rule to determine *WA* and *WW*. Therefore, extended rules to determine *WA* and *WW* are needed. As the third future work, we will propose natural language processing (NLP) methods to extract a criminal' age (*CA*) and weapon (*CW*) from the ERD. This paper showed a keyword matching method to extract the *CA* and *CW* from the simple ERD, but the keyword matching method is not appropriate to extract *CA* and *CW* from complex ERD. Thus, NLP based *CA* and *CW* extractions are needed. These three future works will be done in our future papers.

ACKNOWLEDGMENT

This research was supported and funded by the Korean National Police Agency. [Project Name: 112 Emergency Dispatch Decision Support System/ Project Number:PR08-03-000-21]

REFERENCES

[1] Statistics of Korean National Police Agency, Retrieved from https://www.police.go.kr/www/open/publice/publice0208.jsp

[2] W. B. Shin et al., "Review of the relationship empirically between police workforce size, crime arrest and crime prevention," Korean Police Studies Review, volume 14, issue 4, pp. 163-182, 2015.

[3] G. Choi, "A study on the use of police force," The Journal of Police Science, volume 7, pp. 114-145, October 2004.

[4] The established rule No. 550 of Korean National Police Agency, Retrieved from https://www.police.go.kr/portal/viewer/skin/doc.html?fn=BBS_201907191103059902.hwp&rs=/portal/viewer/201908

[5] Z. Friend, "Predictive policing: using technology to reduce crime," FBI Law Enforcement Bulletin, February 8, 2018.

[6] H. Park et al., "Emergency Dispatch Support System for Police Officers in Incident Scenes," 2021 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1542-1544, October 2021.

[7] H. Park et al., "Crime Response System based on Danger Degree Calculations of Crime Scenes," 2022 International Conference on Information and Communication Technology Convergence (ICTC), pp. 2221-2224, October 2022.

[8] Korean BERT pre-trained cased (KoBERT) - GitHub, Retrieved from https://github.com/SKTBrain/KoBERT