

지휘통제 워크플로우 지원 추천 시스템 연구[☆]

A Study of Recommendation Systems for Supporting Command and Control (C2) Workflow

박 규 동¹ 전 기 윤¹ 손 미 애² 김 종 모*²
Gyudong Park Gi-Yoon Jeon Mye Sohn Jongmo Kim

요 약

정보 통신 및 인공지능 기술의 발전은 우리 군의 지휘통제체계의 지능화를 요구하며, 이를 달성하기 위해 다양한 시도가 이루어지고 있다. 본 논문은 특히, 지휘통제 워크플로우에서 활용 가능한 정보의 양이 폭발적으로 증가함에 따라 지휘통제체계 사용자에게 제공되는 정보 중 수행 업무에 가장 핵심적인 정보를 제공할 수 있는 협업 필터링(Collaborative Filtering, CF) 및 추천 시스템(Recommendation System, RS)에 주목한다. 군 지휘통제체계에서 정보의 필터링을 수행하는 RS는 가장 우선 설명 가능한 추천을 수행하여야 하며, 그 다음 지휘관들이 임무를 수행하는 다양한 상황을 고려한 추천이 수행되어야 한다. 본 논문에서는 지휘통제 워크플로우를 지원하기 위하여 정보를 선택적으로 추천하는 contextual pre-filtering CARS 프레임워크를 제안한다. 제안된 프레임워크는 1) 지휘결심자의 상황 및 관계에 기반하여 데이터를 사전에 필터링하는 contextual pre-filtering, 2) CF의 취약한 데이터 희소성 문제를 극복하기 위한 피쳐 선택, 3) 피쳐 간의 디스턴스를 사용자의 유사도 산출에 활용한 CF, 및 4) 사용자의 선호를 반영하기 위한 규칙 기반 포스트 필터링의 4 단계로 구성되어 있다. 본 연구의 우수성을 평가하기 위해서 상용 수준의 실험 데이터셋 2종에 대해 기존 CF 방법의 다양한 디스턴스 방법을 적용하여 비교 실험하였다. 비교 실험 결과 제안된 프레임워크가 3가지 평가지표(MAE, MSE, MSLE) 측면에서 우수함을 나타내었다.

☞ 주제어 : 지휘통제 시스템, 피쳐 선택, 협업 필터링, 상황 인식 기반 추천 시스템

ABSTRACT

The development of information communication and artificial intelligence technology requires the intelligent command and control (C2) system for Korean military, and various studies are attempted to achieve it. In particular, as a volume of information in the C2 workflow increases exponentially, this study pays attention to the collaborative filtering (CF) and recommendation systems (RS) that can provide the essential information for the users of the C2 system has been developed. The RS performing information filtering in the C2 system should provide an explanatory recommendation and consider the context of the tasks and users. In this paper, we propose a contextual pre-filtering CARS framework that recommends information in the C2 workflow. The proposed framework consists of four components: 1) contextual pre-filtering that filters data in advance based on the context and relationship of the users, 2) feature selection to overcome the data sparseness that is a weak point for the CF, 3) the proposed CF with the features distances between the users used to calculate user similarity, and 4) rule-based post filtering to reflect user preferences. In order to evaluate the superiority of this study, various distance methods of the existing CF method were compared to the proposed framework with two experimental datasets in real-world. As a result of comparative experiments, it was shown that the proposed framework was superior in terms of MAE, MSE, and MSLE.

☞ keyword: Command and Control(C2) system, Feature selection, Collaborative Filtering, Context-aware Recommendation System

1. 서 론

정보 통신 및 인공지능 기술의 비약적인 발전은 우리

군의 지휘통제체계에 대한 지능화와 효율성 향상을 요구하고 있다[1, 2]. 이러한 요구에 부응하기 위해 다양한 시도가 이루어지고 있음에도 불구하고, 지휘통제체계의 지능화 수준은 지루하고 반복적인 업무를 대신하는 수준에 머물고 있다. 인간을 완전히 배제한 지능화의 추진이 불가능한 지휘통제 업무의 특성상, 인간을 배제하기 보다는 지휘관 및 참모의 신속정확한 지휘결심을 지원할 수 있는 수준의 지능화를 추진하는 것이 필요하다[3, 4].

군의 지휘통제체계는 네트워크와 컴퓨터를 활용해 주어진 임무를 완수하는 데 다양한 작전 절차 또는 지휘통

1 Agency for Defense Development, Seoul, Korea

2 Dept. of Industrial Engineering, Sungkyunkwan University, Suwon, Korea

* Corresponding author: (dignityc@gmail.com)

[Received 12 October 2021, Reviewed 20 October 2021(R2 2 December 2021), Accepted 7 December 2021]

☆ 본 연구는 국방과학연구소의 국방 지휘통제 통합-연동 기반기술 특화연구실 과제의 지원을 받아 수행되었습니다(UE201114ED).

제 워크플로우를 지원하는 것을 주 목적으로 한다. 이때 지휘통제 워크플로우란 지휘통제 절차와 관련된 작업, 관련 조직이나 사람, 요구되는 입출력 정보 및 정보의 흐름으로 정의된다. 그러나 최근 센서와 네트워크 기술의 발전으로 지휘통제 워크플로우에서 활용 가능한 정보의 양이 폭발적으로 증가하면서 신속정확한 지휘통제가 어려워지고 있다[5, 6]. 이에 본 논문에서는 지휘통제 워크플로우에서 지휘결심자에게 제공되는 정보를 필터링할 수 있는 지능화 방법 제안함으로써 신속정확한 지휘통제를 달성하고자 한다.

정보의 홍수 속에서 사용자들의 정보처리 부담을 획기적으로 줄여줄 수 있는 방법으로 등장한 것이 추천시스템(recommender system, RS)이다[7]. 전자상거래 분야에서 널리 활용되고 있는 RS는 상용 도메인을 넘어 국방 분야에서도 널리 활용되고 있다[8]. Gadeppally et al은 RS를 사용자를 위해 정보를 선택적으로 필터링하는 것으로 정의하고 사이버 공격 분야에 적용한 바 있으며[9], 미 육군(Army Research Laboratory, ARL)에서도 협업 필터링(Collaborative Filtering, CF) 기반의 추천 알고리즘을 개발한 바 있다[10].

군 지휘통제체계에서 정보의 필터링을 수행하는 RS 개발 시 다음과 같은 요소를 고려해야 한다. 첫째, 지휘통제 업무의 특성상 정보를 추천한 이유를 명확히 설명할 수 있어야 한다. 이는 현재 RS에서 널리 활용되고 있는 모델 기반의 CF 추천이 아니라 메모리 기반의 CF 추천이 수행되어야 함을 의미한다. 메모리 기반 CF 추천은 추천에 사용할 입력 데이터로 활용할 수 있는 요인(예를 들어, 사용자, 아이템, 상황, 관계 등) 간의 관계를 직접적으로 도출하는 것이 가능하다. 특히 특정 사용자(군에 소속되거나 군 관련 임무를 수행하는 인원)와 아이템(상황도, 각종 화면, 계시관, 자료실 등의 기능 및 정보) 간의 관계를 명시적으로 설명하면서도 도메인 지식을 요구하지 않는 CF가 가장 적합한 추천 방법이다. 둘째, 지휘관들에게는 임무를 수행하는 데 필요한 정보 및 사용자들의 관계(예를 들어, 계급, 직책 등)를 고려한 정보가 제공되어야 한다. 이는 사용자(지휘관) × 아이템(정보)을 사용하는 전통적인 CF가 아니라 사용자×아이템×상황×관계를 종합적으로 고려할 수 있는 확장된 CF여야 함을 의미한다. 확장된 CF는 상황 요인이 데이터 차원에서 전처리된 contextual pre-filtering CARS (context-aware recommender system)라 할 수 있다.

본 논문에서 제안하는 지휘통제 워크플로우에서 정보를 선택적으로 추천하는 contextual pre-filtering CARS 프

레이워크의 구성요소는 다음과 같다.

- contextual pre-filtering: 상황과 관계를 반영할 수 있는 입력데이터 생성 과정으로 사용자, 상황 및 관계를 고려하여 관계가 없는 데이터를 사전에 필터링함으로써 CF가 계산해야 할 데이터의 규모를 축소한다
- 피쳐 선택: CF가 가진 고질적인 문제는 데이터 희소성(data sparsity)이다. 본 논문에서 제안하는 RS 프레임워크의 경우 사용자×아이템×상황×관계 데이터를 사용하기 때문에 데이터 희소성 문제는 전통적인 CF 보다 훨씬 더 심각하다. 왜냐하면 데이터 희소성 문제는 데이터의 차원이 높아질수록 심각해지기 때문이다. 이를 완화하기 위해서는 피쳐 선택을 수행해야 한다.
- 메모리 기반의 CF: 지휘관(사용자)과 정보(아이템) 간의 관계를 명시적으로 설명할 목적으로 수행한다.
- 규칙 기반 포스트 필터링: 메모리 기반의 CF 결과를 대상으로 사전에 정의되어 있는 도메인 지식이나 사용자의 선호 등을 반영해 포스트 필터링을 수행한다. 도메인 지식이나 사용자의 선호는 규칙 형식으로 표현하며 사전에 식별되어 있다고 가정한다. 규칙 기반의 포스트 필터링을 통해 사용자가 원하는 정보만을 추천하는 것이 가능해진다.

본 논문의 구성은 다음과 같다. 2장에서 기존의 연구들을 살펴보고, 3장에서 본 논문에서 제안한 방법의 프레임워크와 각 모듈에 대해 설명한다. 4장에서 실험과 그 결과에 대해 설명하고 5장에서 결론 및 향후 연구에 대해 논의한다.

2. 관련 연구

전통적으로 추천 시스템은 콘텐츠(Content) 기반 추천 시스템과 협업 필터링(Collaborative) 기반 추천 시스템, 및 이를 혼합한 하이브리드 추천 시스템으로 구분된다. 그 중에서도 협업 필터링은 데이터를 직접 계산하는 산출식을 활용하는지 혹은 머신 러닝의 모델에 의존하는지에 따라, 메모리 기반 및 모델 기반으로 분류된다. 또한, 이러한 추천 시스템의 입력값에 사용자 및 아이템에 더하여 상황(Context) 요소까지 추가적으로 고려할 경우 상황 인식 추천 시스템으로(CARS) 분류된다[11, 12].

CARS의 핵심은 사용자의 아이템 구매에 영향을 미치는 상황의 패턴을 데이터로부터 정확하게 발견하는 것이다. 따라서, 상황의 패턴을 발견하는 방법, 즉 상황 모델

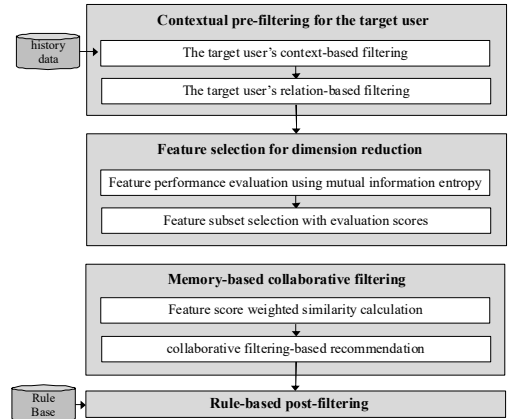
링하는 과정을 모델 의존적으로 하는지에 따라서 연구들을 구분할 수 있다[13]. 첫 번째로, 모델에 기반한 상황 모델링은 *tensor factorization* 및 *Convolutional matrix factorization*과 같은 머신 러닝 모델을 활용하여 상황을 모델링한다. 이 방법은, 머신 러닝의 우수한 예측 성능을 활용하기 때문에 복잡한 상황 정보로부터 잠재적인 (latent) 상황까지 정확하게 모델링할 수 있다는 장점이 있다[14-20]. 그러나, 설명 가능한 추천이 어렵다는 점과 충분한 일반화 성능을 위해서 고품질의 레이블 데이터를 요구하는 한계가 있다.

반면 모델에 비의존적으로 상황 모델링을 수행하는 연구들은 다양한 방법들을 시도해왔다. 그래프 기반의 상황 모델링 방법은 상황 정보와 유저의 선호를 특정 데이터의 피쳐나 차원에 귀속되지 않고 자유롭게 표현하기 위해 활용되었다[21]. 피쳐 기반의 상황 모델링 방법은 상황을 구성하는데 핵심이 되는 피쳐와 이의 가중치를 발견하여, 유의미한 상황이 도출되는 조건을 발견한다[22]. 이러한 방법들은 특정 도메인의 지식을 활발하게 활용하기 때문에 보다 정확하고 일반화된 상황 모델링이 가능하나 전문가의 지식이 필요하기 때문에 상당한 비용을 야기한다.

또한, 지식 기반 추천 시스템(KBRs)은 도메인 지식을 활용하여 데이터 처리의 부담 없이 고품질의 추천을 수행한다[23]. 왜냐하면 KBRs는 대규모 데이터 셋의 필요, 콜드 스타트 문제 등으로부터 어느 정도 자유롭기 때문이다. 그러나, 지식을 구축하는 자체가 상당히 제약이 되기 때문에 확장 가능성이 CARS나 기타 데이터 활용 추천 시스템에 비해 현저히 낮다.

3. 지휘통제 워크플로우를 위한 CF 기반 contextual pre-filtering CARS 프레임워크

본 논문에서는 지휘통제 워크플로우에서 정보를 선택적으로 추천할 수 있는 CF 기반의 contextual pre-filtering CARS 프레임워크를 제안한다. 그림 1에 도식화한 바와 같이 제안 프레임워크는 4가지 요소로 구성되어 있다.



(그림 1) 지휘통제 워크플로우를 위한 CF 기반 contextual pre-filtering CARS 프레임워크

(Figure 1) A CF-based Contextual pre-filtering CARS framework for C2 workflow

3.1 타겟 유저를 위한 Contextual pre-filtering

전술한 바와 같이, 지휘통제 워크플로우를 위한 CARS 프레임워크는 사용자×아이템×상황×관계를 고려해 정보를 선택적으로 필터링하는 추천시스템이다. 그러나 CF를 적용하기 위해서는 사용자×아이템×상황×관계의 고차원 데이터를 2차원 행렬로 나타내야 한다. 이를 위해, 일련의 피쳐와 이의 데이터로 조합된 희소 표현(sparse representation)을 활용한다. 지휘통제 워크플로우를 위한 CARS 프레임워크에서 사용할 소스 데이터는 다음과 같다.

정의 1. i^{th} 소스 데이터(x_i) x_i 는 특정 지휘관이 임의의 정보를 소비한 상황 및 관계를 기록한 과거 데이터로 다음과 같이 표현된다.

$$x_i = \{u_i, i_i, C_i, L_i, r_i\} x_i \in X, \forall i \quad (1)$$

이 때, u_i 는 i^{th} 데이터의 사용 주체인 사용자의 ID ($u_i \in U$), i_i 는 u_i 가 소비한 아이템의 ID ($i_i \in I$), C_i 는 i^{th} 데이터가 생성된 상황을 나타내는 집합으로 날짜, 요일, 시간대와 같은 시간 정보 집합(t_i)과 위도, 경도, Point-of-Interest 태그와 같은 공간 정보 집합(l_i)을 포함한다 ($C_i = \{t_i, l_i\}$, $C_i \in C$). 마지막으로 L_i 는 u_i 가 i_i 를 소비했을 때 상호작용했던 다른 사용자와의 관계 정보로 직속상관, 협업 관계와 같은 관계의 성질을 태그로 나타

내는 값(rel_i)과 이의 대상인 사용자 ID 나타내는 값 (u_i^{obj})의 두 원소로 구성된다($L_i = \{(rel_i, u_i^{obj})\}$). r_i 은 i_i 에 대한 레이팅 값을 의미한다.

2차원 희소 표현 데이터는 상황과 관계를 기반으로 사전 필터링을 수행한다. 사전 필터링의 목적은 CF의 부담을 줄이기 위해서이다. 이를 위해, 먼저 정보 제공의 대상 즉 타겟 사용자(u^t)와 관련된 데이터를 X 로부터 추출한다. 임의의 타겟 사용자 $u^t (u^t \in U)$ 가 소비한 데이터는 다음과 같다.

$$d(u^t) = \{x_j | u_j = u^t, \forall j, j \leq i\} \quad (2)$$

3.1.1 상황기반 필터링

소스 데이터 X 부터 임의의 사용자 u^t 의 지휘결심이 필요한 상황 (C^t)과 연관성이 없는 데이터를 필터링한다. 이때 지휘결심이 필요한 상황 (C^t)는 다음과 같이 정의된다.

$$C^t = \{C_j | C_j \in d(u^t), \forall j, j \leq i\} \quad (3)$$

u^t 의 상황인 C^t 과 유사하지 않은 상황을 필터링하기 위해서, 타겟 상황과 임의의 상황 간의 유사도를 산출한다. 일반적으로 상황 데이터는 숫자형 및 범주형 데이터의 혼합이기 때문에 일반적인 거리 함수를 이용해 유사도를 도출하는 것이 어렵다. 따라서, 본 논문에서는 데이터의 형식에 비의존적인 정규화된 상호 의존 정보(mutual information)를 활용해 유사도를 산출하고자 한다. C^t 와 임의의 상황(C_k)간의 유사도를 다음과 같이 산출한다.

$$sim(C_k, C^t) = \frac{I(C_k, C^t)}{(\min(H(C_k), H(C^t)))} \quad (4)$$

이때, $I(C_k, C^t)$ 는 C_k 와 C^t 간의 상호 의존 정보이고 $H(C_k)$ 및 $H(C^t)$ 는 C_k 와 C^t 의 정보 엔트로피이다 ($k \neq j, k \in i$).

$sim(C_k, C^t)$ 가 정해진 임계치 이하인 x_i 를 X 에서 필터링함으로써 상황과 관련이 있는 x_k 만을 획득한다. 타겟 사용자와 상황적으로 관련이 있는 x_k 는 다음과 같다.

$$d(C^t) = \{x_k | sim(C_k, C^t) \geq \pm \delta\} \quad (5)$$

3.1.2 관계 기반 필터링

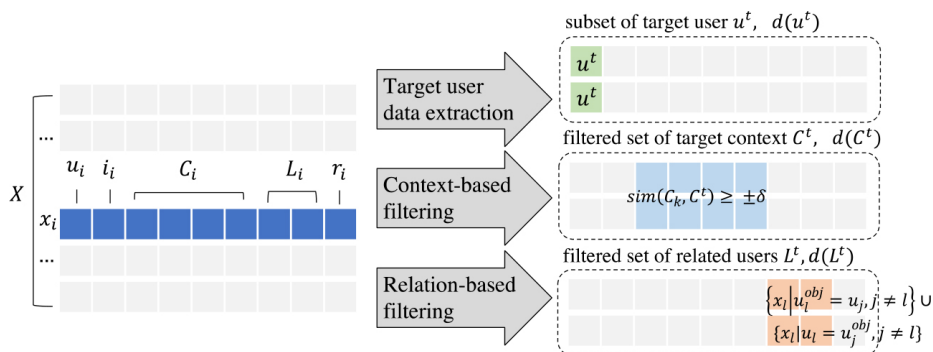
소스 데이터 X 부터 지휘통제 절차상 u^t 와 관계가 없는 사용자의 데이터를 필터링한다. 이때 관계는 u^t 에게 정보를 송신하는 송신자로서의 관계에 있는 사용자와 u^t 로부터 정보를 수신하는 수신자 관계에 있는 사용자를 포함한다. u^t 관계를 갖는 사용자의 기록데이터 x_i 는 다음과 같은 기준으로 선별한다.

$$d(L^t) = \{x_l | u_i^{obj} = u_j, j \neq l\} \cup \{x_l | u_i = u_j^{obj}, j \neq l\} \quad (6)$$

이상의 과정을 거쳐 필터링된 데이터는 다음과 같이 표현된다.

$$D_i = d(u^t) \cup d(C^t) \cup d(L^t) \quad (7)$$

데이터를 필터링하는 과정과 생성되는 데이터의 구조를 개념적으로 도식화하면 그림 2와 같다.



(그림 2) 타겟 사용자의 상황 및 관계 기반의 contextual pre-filtering
(Figure 2) Context and relation-based contextual pre-filtering for the target user

3.2 차원 축소를 위한 피쳐 선택

그림 2에 도식화한 바와 같이, 상황과 관계를 고려해 필터링을 수행한 데이터셋 D_t 는 여전히 null 값이 다수인 데이터 희소성이 심각한 데이터이다. 데이터 희소성은 사용자와 아이템 간의 관계에 대한 정확한 이해를 어렵게 하는 CF의 고질적인 문제이다[24]. 이러한 문제를 해결하는 방법으로 D_t 의 차원을 축소하는 피쳐 선택 기법이 널리 활용된다 [25]. 피쳐 선택은 독립적인 피쳐 평가 함수를 활용하는 필터 기반 방법과 임의의 피쳐 서브셋에 대한 학습된 분류 결과에 따라서 최적의 서브셋을 탐색하는 래퍼 기반 방법 또는 이들을 하이브리드한 방법[26]을 통해 수행한다. 본 논문에서 활용하고자 하는 데이터와 같이 숫자형 및 범주형의 데이터가 혼합된 경우 피쳐 평가 함수의 편향성이 발생하여 필터 방법의 성능이 현저히 떨어지기 때문에 본 논문에서는 래퍼 기반의 피쳐 선택 방법을 활용해 차원을 축소한다.

래퍼 기반의 피쳐 선택 절차는 다음과 같이 요약된다. 첫째, 불필요한 피쳐 서브셋 평가를 피하기 위해 정보 엔트로피를 활용해 개별 피쳐의 단순 성능을 평가하고 정렬한다. 둘째, 성능이 가장 우수한 피쳐부터 순차적으로 임시 피쳐 서브셋을 생성하고 이를 입력 데이터로 활용해 분류 모델을 학습시킨다. 셋째, 전 단계와의 분류 성능을 비교한다. 전 단계보다 분류 성능이 좋다면 현 단계의 임시 피쳐 서브셋에 포함된 피쳐를 최종 피쳐 서브셋에 포함시킨다. 마지막으로 임시 피쳐 서브셋에 새로운 피쳐를 추가해도 분류 성능이 향상되지 않을 때까지 (2)~(3)의 과정을 반복한다. 이상의 과정을 상세히 표현하면 알고리즘 1과 같다. 결과로 축소된 차원의 데이터셋 D_t^- 를 획득한다.

3.3 메모리 기반 협업 필터링

전통적인 CF는 사용자가 아이템에 대해 평가한 레이팅으로부터 사용자 혹은 아이템에 기반한 유사도를 산출하여 아이템을 추천한다. 그러나 지휘통제에 필요한 정보나 기능은 임무상황과 사용자의 관계에 따라 달라질 수 있다. 그러므로 정확한 정보나 기능의 추천을 위해서는 상황과 관계 정보도 함께 고려한 사용자 기반의 유사도 계산이 필요하다. 이를 위해 본 논문에서는 앞서 유의 미한 피쳐로 축소된 데이터셋 D_t^- 에서 두 사용자의 상황과 관계의 차이를 가중치로 활용하는 피어슨 유사도를 제안한다. 일반적으로 피어슨 유사도를 활용하여 두 사용자의 유사도는 다음과 같이 계산한다.

Algorithm 1. Wrapper-based feature subset selection

Input: D_t
Output: Reduced set D_t^- , Selected Features FS

```

1: Features = []
2: For j in range( $D_t$ .shape(1)):
3:   vectors = []
4:   For i in range( $D_t$ .shape(0)):
5:     Features.append(vectors)
6: Evaluator = Evaluation('information_entropy')
7: SortedSet = Evaluator(Features).sorted(Descending)
8: Temp = SortedSet.leftpop(1)
9: Searching = True
10: While Searching:
11:   For idx in range(len(Features)):
12:     cnt = Temp+idx
13:     Acc_cnt = wrapper(cnt)
14:     Acc_temp = wrapper(temp)
15:     If acc_cnt > acc_temp:
16:       Temp = cnt.copy()
17:       Features.pop(index=idx)
18:     break
19:   If acc_cnt < acc_temp:
20:     Searching=False
21:  $D_t^- = D_t[:,Temp]$ 
22: FS = Temp.copy()

```

$$Sim(u, u') = \frac{\sum_{u, u' \in U} (R_u - \overline{R_u})(R_{u'} - \overline{R_{u'}})}{\sqrt{\sum_{u, u' \in U} (R_u - \overline{R_u})^2 \sum_{u \in U} (R_u - \overline{R_u})^2}} \quad (8)$$

이때, R_u 은 임의의 사용자 u 가 모든 아이템에 대해 평가한 레이팅 벡터이다. 또한, $\overline{R_u}$ 은 R_u 의 평균값이다.

축소된 데이터셋 D_t^- 로부터 두 사용자의 상황과 관계의 차이를 산출하기 위해 임의의 사용자 u 에 대한 D_t^- 의 부분집합은 다음과 같이 정의된다.

$$w_u = \{x_u | u_u = u, x_u \in D_t^-\} \quad (9)$$

이때, X_u 의 값은 max-scaling 되어 있음을 가정한다. 식 (8) 및 (9)에 기반하여 두 사용자 간의 피쳐 값의 차

이가 가중치로 부여된 피어슨 유사도는 다음과 같다.

$$\text{weightd-Sim}(u, u') = \frac{\sum_{u,f} w_u w_{u'} (R_{u-f} - \overline{R_u})(R_{u'-f} - \overline{R_{u'}})}{\sqrt{\sum_{u,f} (W_u)(R_{u-f} - \overline{R_u})^2} \sqrt{\sum_{u',f} (W_{u'})(R_{u'-f} - \overline{R_{u'}})^2}} \quad (10)$$

사용자 간의 유사도를 산출할 때 단순히 레이팅의 차이만 계산하는 것이 아니라 피쳐의 차이를 계산함으로써 사용자의 상황 및 관계의 차이도 고려된 복합적인 유사도를 산출할 수 있다. 식 (10)으로 산출된 사용자 간의 유사도에 기반하여 CF를 수행한 후, 추천된 아이템 리스트를 획득한다.

3.4 규칙 기반 포스트 필터링

추천된 아이템 리스트는 비록 사용자의 과거의 소비 데이터에 기반한 아이템들의 집합이지만, 사용자의 선호를 직접적으로 반영하지 않았기 때문에 사용자가 원치 않는 아이템이 추천 리스트에 포함될 수 있다. 사용자의 선호를 반영함과 동시에 추천된 아이템 리스트의 정확도를 향상시키기 위해, 본 논문에서는 최종적으로 규칙 기반의 포스트 필터링을 수행한다. 도메인 지식 혹은 사용자의 선호는 규칙 형식으로 표현되어 있는 규칙 베이스가 사전에 구축되어 있음을 가정한다. 지휘통제에서 활용할 수 있는 규칙의 예시는 다음과 같다.

위치 요소와 관련하여 다음과 같은 규칙을 활용한다.

[R1] 상급부대에서 제공된 정보의 중요도가 하급 부대에서 제공보다 중요하다.

[R1-1] 사용자 소속 부대의 관심 지역 외 정보는 관심 지역 내 정보보다 제공 우선 순위가 낮다.

많은 지휘통제 워크플로우들이 주기에 따라 반복적으로 수행된다. 하지만 그 주기의 시작 시점이 때로는 달라질 수 있다. 대신 순서는 거의 변하지 않는다. 예를 들어 다음과 같은 유형의 규칙을 활용할 수 있다.

[R2] 화력 임무 워크플로우는 표적평가, 무기할당, 타격결과관리 순으로 수행된다.

상기 제안된 규칙들은 기능 타당성을 설명하기 위한 예시용 규칙이며, 실제 규칙 및 지식은 이와 다를 수 있다. 식별된 규칙의 입력 변수에 입력값을 사용자의 기록 데이터로부터 발견하고 이를 적용하여 규칙 기반 포스트 필터링을 수행한다.

4. 실험 및 평가

보안상의 이유로 군 관련 실험 데이터를 확보하는 것이 어렵기 때문에 본 논문에서는 실제 상용 데이터셋을 군 환경에 맞게 변환하여 제안한 방법론의 우수성을 평가한다. 상용 데이터셋 중 실제 상황 데이터를 포함하며 군 환경과 같이 유저의 위치 변화가 빈번하고 다양한 유형의 아이템을 소비하는 데이터셋을 선택하였다. 또한, 적합한 군 환경을 모사하기 위해 상용 데이터셋에서 가장 유사한 군 환경을 식별하여 기존 분포에 큰 영향을 주지 않는 선에서 데이터셋을 재가공 하였다.

첫 번째로 활용한 상용 데이터셋은 차에서 청취한 음악에 대한 InCarMusic 데이터셋이다. 이 데이터를 군 환경에 비슷하게 모사하기 위해 이동 중에 제공받는 메시지 데이터셋 MobileMsg으로 변환한다. 운행에 관한 환경 정보에서 날씨 및 위치 정보를 활용하였으며, 또한 음악의 장르는 메시지의 유형으로 변환하였다. 이 과정에서 '분위기' 혹은 '아티스트명'과 같이 군 환경과 관련 없는 데이터는 제거하였다. 두 번째로 활용한 상용 데이터셋은 Restaurant 데이터셋으로, 이를 군 환경과 비슷하게 모사하기 위해 순찰한 장소에 대한 트랙 보고 데이터셋 TrackReport로 변환하였다. 이때, 장소와 관련된 위치 및 지오태그 정보를 포함하였으며 유저 프로필과 관련된 정보는 직업 및 함께 방문한 유저 정보를 포함하여 변환하였다. 각 변환한 실험 데이터 셋의 특징에 대한 요약은 다음 표 1과 같다.

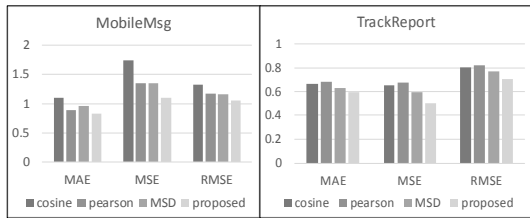
(표 1) 실험 데이터 셋에 대한 요약

(Table 1) The summary of experimental data sets

Properties	MobileMsg	TrackReport
#Records	2,593	1,161
#Users	42	138
#Items	139	130
#Features	4	36
Types of context	Location, Sentiment	Time, Location, User profile
Range of ratings	1 - 5	0 - 2
Data types	Categorical	Mixed
Ratio of Nulls	78.6492%	8.9625%

위 두 실험 데이터 셋과 함께 비교 실험할 방법들로 사용자 기반의 CF를 활용하였으며 사용자의 유사도를 산출하는 거리 함수로 코사인(Cosine) 유사도, 피어슨

(Pearson) 유사도, 및 Mean Squared Difference (MSD)를 활용하였다. 유사도의 경우 디스턴스를 나타내기 위해 (1-유사도)를 적용하였다. 평가 지표로는 Mean Absolute Error(MAE), Mean Squared Error(MSE) 및 Root Mean Square(RMSE)를 활용하였다.



(그림 3) 사용자 기반 CF에서 다른 디스턴스를 활용한 방법과 성능 비교

(Figure 3) Performance comparison with different distance metrics in user-based CF

MAE는 실제 값과 예측 값의 차이의 절대값의 평균이다. MAE는 차이를 평면적으로 보여주기 때문에 에러의 산포도나 경향성을 반영하기 어렵다. 반면, MSE는 실제 값과 예측 값의 차이를 제곱한 값의 평균이기 때문에 차이가 크면 클 수록 값이 가파르게 상승한다. 흔히 모델의 안정성을 평가할 때 활용한다. MSE는 에러를 제공하기 때문에 필요 이상으로 이상치(아웃라이어)의 에러를 크게 산출한다. RMSE는 주로 정밀도를 가늠하는 데 활용된다. 모든 실험 결과는 5-fold cross validation의 평균값이다. 실험 결과의 상세 값은 표 2와 같으며 비교 그래프는 그림 3과 같이 나타내었다.

두 데이터셋의 에러가 다르게 나오는 이유는 활용한 레이팅 값의 범위가 서로 차이가 있기 때문이다. 5점 척

도를 활용하는 MobileMsg 데이터셋이 더 높은 에러를 나타냈지만 3점 척도를 활용하는 TrackReport와 비교하면 레이팅을 예측하기 더 어려운 점을 고려하였을 때 상대적으로 지표가 좋게 나왔음을 알 수 있다. 이는 데이터셋의 분포 차이 때문인데, MobileMsg는 사용자-아이템 간의 상호작용 데이터가 많이 존재하는 dense인데 비해 TrackReport는 이러한 상호작용 데이터가 현저하게 부족한 sparse set이다. 다시 말해, 피쳐의 Null값은 InCarMusic이 적었으나 데이터의 희소성은 TrackReport 데이터가 훨씬 높은 수준이었다.

다른 방법과 비교하였을 때, 모든 지표에서 제안된 방법이 최소의 에러를 나타내는 것을 볼 수 있다. Cosine 유사도를 활용한 추천은 MobileMsg 데이터셋에 대해서 다른 디스턴스에 비해 상당히 높은 에러를 나타내었다. 이는, 5점 척도에서 Cosine 유사도가 레이팅의 scale을 제대로 반영하지 못하기 때문이다. TrackReport에서는 척도의 범위가 더 좁아서 지표 간 에러의 큰 차이가 없었다. 본 연구에서 제안한 방법은 피어슨 유사도에 기반하였으나 피쳐의 값을 가중치로 활용하였기 때문에 TrackReport와 같이 이상치가 많이 존재하는 데이터에 대해서도 충분히 보정되었음을 알 수 있다. 결과적으로, 두 데이터셋 및 3개의 평가 지표에서 본 연구에서 제안한 방법이 안정적이고 높은 정확도로 추천을 수행함을 나타내었다.

5. 결론 및 추후 연구

본 연구는 한국군 지휘통제체계를 위한 추천 시스템에 대한 첫 번째 구현 또는 구현 시도로서 의미가 있다. 본 연구에서는 지휘통제 워크플로우에서 활용 가능한 정보의 양이 폭발적으로 증가함에 따라, 지휘통제체계 사용자에게 효과적으로 필요한 정보를 추천할 수 있는 협업 필터링 및 추천 시스템에 관한 연구를 수행하였다. 기

(표 2) 실험 데이터셋에 대해 다양한 디스턴스를 활용한 사용자 기반 CF의 성능 평가

(Table 2) Evaluation results of User-based CF with different distances on experimental datasets

Methods		User-based CF			Proposed method
		Cosine similarity	Pearson coefficient	Mean Squared Difference (MSD)	
MobileMsg	MAE	1.0998	0.8888	0.9512	0.8289
	MSE	1.7424	1.3487	1.3478	1.1032
	MSLE	1.3195	1.1612	1.1603	1.0503
TrackReport	MAE	0.6638	0.6836	0.6303	0.5942
	MSE	0.6516	0.6786	0.5946	0.5003
	MSLE	0.8009	0.8236	0.7709	0.7069

존의 협업 필터링 방법과 다르게 CARS 및 KBRs의 개념을 접목하여 협업 필터링의 수행의 대상이 되는 데이터셋을 contextual pre-filtering 및 rule-based post-filtering하였다. 또한, CF의 성능을 저해하는 데이터의 희소성을 막기 위해 다양한 피쳐 중에서 추천에 가장 큰 영향을 주는 피쳐를 선택할 수 있는 피쳐 선택 기법을 제안하였다. 최종적으로, 본 논문의 우수성을 평가하기 위해 기존 사용자 기반의 CF에서 다양한 디스턴스를 적용하여 비교 실험하였고, 그 결과 MAE, MSE, 및 MSLE의 세 가지 지표에서 모두 본 연구가 제안한 프레임워크가 우수한 성능을 나타내었다.

우리는 본 연구의 결과를 후속 연구를 통해 지속적으로 개선 및 발전시킬 예정이다. 예를 들어 사용자, 상황, 및 관계에 따라 소스 데이터의 구조가 달라질 수 있고 시간에 따라 역동적으로 변화할 수 있어, 소스 데이터의 복잡도 또는 차원이 모수 통계를 기반으로 하는 제안 프레임워크의 처리 능력을 넘어설 수도 있다. 이러한 한계를 극복하기 위해, 추후에 우리는 그래프 기반의 소스 데이터 표현 및 그래프 인공지능망을 활용하는 방법에 대해 연구할 것이다. 그래프 기반의 인공지능망을 활용하면 더 이상 특정 소스 데이터의 특성에 의존하지 않더라도 자유로운 데이터의 표현과 복잡한 사용자, 상황 및 관계를 이해할 수 있는 최적의 초평면(hyper-plane)을 발견할 수 있을 것이다. 이를 포함하여 우리는 궁극적으로 한국군 지휘통제체계의 추천 시스템의 실제 개발에 활용 가능한 수준의 기술과 기능을 개발하고자 한다.

References

- [1] Schubert, J., et al., "Artificial intelligence for decision support in command and control systems," in Proc. of 23rd International Command and Control Research & Technology Symposium Multi-Domain C., 2018.
https://www.foi.se/download/18.41db20b3168815026e010/1548412090368/Artificial-intelligence-decision_FOI-S-5904-SE.pdf
- [2] Wang, F., "Technology Framework of the Intelligent Command and Control System," in IOP Conference Series: Materials Science and Engineering, IOP Publishing, 2019.
<https://doi.org/10.1088/1757-899x/677/4/042099>
- [3] Lingel, S., et al., "Joint All Domain Command and Control for Modern Warfare: An Analytic Framework for Identifying and Developing Artificial Intelligence Applications," RAND Corporation Santa Monica United States, 2020. <https://doi.org/10.7249/tr4408.1>
- [4] Robinson, R.M., et al., "Degree of automation in command and control decision support systems," in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2016.
<https://doi.org/10.1109/SMC.2016.7844402>
- [5] Zocco, A., et al., "Touchless interaction for command and control in military operations," in International Conference on Augmented and Virtual Reality, Springer, 2015.
https://doi.org/10.1007/978-3-319-22888-4_32
- [6] Marusich, L.R., et al., "Effects of information availability on command-and-control decision making: performance, trust, and situation awareness," Human factors, 58(2): p. 301-321, 2016.
<https://doi.org/10.1177%2F0018720815619515>
- [7] Guo, Y., M. Wang, and X. Li, "Application of an improved Apriori algorithm in a mobile e-commerce recommendation system," Industrial Management & Data Systems, 2017.
<https://doi.org/10.1108/IMDS-03-2016-0094>
- [8] Chen, C.-s., et al., "Research on Content Recommendation System of Army Equipment Information Network," in 2020 International Conference on Computer Science and Management Technology (ICCSMT), IEEE, 2020.
<https://doi.org/10.1109/ICCSMT51754.2020.00075>
- [9] Gadepally, V.N., et al., "Recommender systems for the department of defense and intelligence community," Lincoln Laboratory Journal, 22(1), 2016.
<https://www.ll.mit.edu/sites/default/files/publication/doc/2019-04/recommender-systems-department-defense-intelligence-gadepally-108929.pdf>
- [10] Yang, C., et al., "Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017.
<https://doi.org/10.1145/3097983.3098094>
- [11] Kulkarni, S. and S.F. Rodd, "Context Aware

- Recommendation Systems: A review of the state of the art techniques," *Computer Science Review*, 37, p. 100255, 2020.
<https://doi.org/10.1016/j.cosrev.2020.100255>
- [12] Adomavicius, G. and A. Tuzhilin, "Context-aware recommender systems," *Recommender systems handbook*, p. 217-253, Springer, 2011.
https://doi.org/10.1007/978-0-387-85820-3_7
- [13] Zheng, Y., B. Mobasher, and R. Burke, "Similarity-based context-aware recommendation," in *International Conference on Web Information Systems Engineering*, Springer, 2015.
https://doi.org/10.1007/978-3-319-26190-4_29
- [14] Unger, M., et al., "Towards latent context-aware recommendation systems," *Knowledge-Based Systems*, 104, p. 165-178, 2016.
<https://doi.org/10.1016/j.knosys.2016.04.020>
- [15] Kim, D., et al., "Convolutional matrix factorization for document context-aware recommendation," in *Proceedings of the 10th ACM conference on recommender systems*, 2016. <https://doi.org/10.1145/2959100.2959165>
- [16] Xin, X., et al., "CFM: Convolutional Factorization Machines for Context-Aware Recommendation," in *IJCAI*, 2019.
<https://www.ijcai.org/proceedings/2019/0545.pdf>
- [17] Chen, H. and J. Li, "Adversarial tensor factorization for context-aware recommendation," in *Proceedings of the 13th ACM Conference on Recommender Systems*, 2019. <https://doi.org/10.1145/3298689.3346987>
- [18] Tao, S., et al., "SVD-CNN: a convolutional neural network model with orthogonal constraints based on SVD for context-aware citation recommendation," *Computational Intelligence and Neuroscience*, 2020.
<https://doi.org/10.1155/2020/5343214>
- [19] Baltrunas, L., B. Ludwig, and F. Ricci, "Matrix factorization techniques for context aware recommendation," in *Proceedings of the fifth ACM conference on Recommender systems*, 2011.
<https://doi.org/10.1145/2043932.2043988>
- [20] Cui, L., et al., "A novel context-aware recommendation algorithm with two-level SVD in social networks," *Future Generation Computer Systems*, 86, p. 1459-1470, 2018.
<https://doi.org/10.1016/j.future.2017.07.017>
- [21] Wu, H., et al., "Context-aware recommendation via graph-based contextual modeling and postfiltering," *International Journal of Distributed Sensor Networks*, 11(8): p. 613612, 2015.
<https://doi.org/10.1155%2F2015%2F613612>
- [22] Zammali, S. and S.B. Yahia, How to select and weight context dimensions conditions for context-aware recommendation? *Expert Systems with Applications*, p. 115176, 2021.
<https://doi.org/10.1016/j.eswa.2021.115176>
- [23] Bouraga, S., et al., "Knowledge-based recommendation systems: A survey," *International Journal of Intelligent Information Technologies (IJIT)*, 10(2), p. 1-19, 2014.
<https://doi.org/10.4018/ijit.2014040101>
- [24] Grčar, M., et al., "Data sparsity issues in the collaborative filtering framework," in *International workshop on knowledge discovery on the web*, Springer, 2005. https://doi.org/10.1007/11891321_4
- [25] Ramezani, M., P. Moradi, and F.A. Tab., "Improve performance of collaborative filtering systems using backward feature selection," in *The 5th Conference on Information and Knowledge Technology*, IEEE, 2013.
<https://doi.org/10.1109/IKT.2013.6620069>
- [26] Kim, J., J. Kang, and M. Sohn, "Ensemble learning-based filter-centric hybrid feature selection framework for high-dimensional imbalanced data," *Knowledge-Based Systems*, 220, p. 106901, 2021.
<https://doi.org/10.1016/j.knosys.2021.106901>

◎ 저 자 소 개 ◎



박 규 동(Gyudong Park)

1994년 홍익대학교 컴퓨터공학과 졸업
1996년 홍익대학교 컴퓨터공학과 석사
2014년 홍익대학교 컴퓨터공학과 박사
1996년 3월~현재 국방과학연구소 연구원
관심분야: C4I, 가상화, 클라우드, 네트워크
E-mail: iobject@add.re.kr



전 기 윤(Gi-Yoon Jeon)

2000년 동국대학교 컴퓨터공학과 졸업
2002년 포항공과대학교 정보통신학과 석사
2021년 동국대학교 컴퓨터공학과 박사
2002년 1월~현재 국방과학연구소 연구원
관심분야: C4I, Human-Robot Interaction, IoT/IoB, 인공지능, Visualization
E-mail: gyjeon@add.re.kr



손 미 애(Mye Sohn)

1985년 성균관대학교 산업공학과(공학사)
1988년 한국과학기술원 산업공학과(공학석사)
2002년 한국과학기술원 경영공학과(공학박사)
2004년 ~ 현재 성균관대학교 산업공학과 교수
관심분야 : 인공지능/전문가시스템, 지식그래프, 시맨틱웹, 온톨로지, IOT, 기계학습, 추천시스템
E-mail : myesohn@skku.edu
[ORCID: 0000-0002-1951-3493]



김 종 모(Mye Sohn)

2014년 성균관대학교 시스템경영공학과 졸업
2021년 성균관대학교 산업공학과 졸업 (석박사통합과정)
2021년 ~ 현재 성균관대학교 산업공학과 박사후연구원
2004년 ~ 현재 성균관대학교 산업공학과 교수
관심분야 : 그래프 마이닝, 지식그래프, 시맨틱웹, 링크드 오픈 데이터, IOT, 기계학습
E-mail : dignityc@gmail.com
[ORCID: 0000-0002-4984-1674]