

Hybrid Offloading Technique Based on Auction Theory and Reinforcement Learning in MEC Industrial IoT Environment

Bae Hyeon Ji[†] · Kim Sung Wook^{††}

ABSTRACT

Industrial Internet of Things (IIoT) is an important factor in increasing production efficiency in industrial sectors, along with data collection, exchange and analysis through large-scale connectivity. However, as traffic increases explosively due to the recent spread of IIoT, an allocation method that can efficiently process traffic is required. In this thesis, I propose a two-stage task offloading decision method to increase successful task throughput in an IIoT environment. In addition, I consider a hybrid offloading system that can offload compute-intensive tasks to a mobile edge computing server via a cellular link or to a nearby IIoT device via a Device to Device (D2D) link. The first stage is to design an incentive mechanism to prevent devices participating in task offloading from acting selfishly and giving difficulties in improving task throughput. Among the mechanism design, McAfee's mechanism is used to control the selfish behavior of the devices that process the task and to increase the overall system throughput. After that, in stage 2, I propose a multi-armed bandit (MAB)-based task offloading decision method in a non-stationary environment by considering the irregular movement of the IIoT device. Experimental results show that the proposed method can obtain better performance in terms of overall system throughput, communication failure rate and regret compared to other existing methods.

Keywords : Mobile Edge Computing(MEC), Device-to-Device(D2D) Offloading, Industrial Internet of Things(IIoT), McAfee's Double Auction, Multi-Armed Bandit(MAB)

MEC 산업용 IoT 환경에서 경매 이론과 강화 학습 기반의 하이브리드 오프로딩 기법

배 현 지[†] · 김 승 옥^{††}

요 약

산업용 IoT는 대규모 연결을 통해 데이터 수집, 교환, 분석과 함께 산업 분야의 생산 효율성 개선에 중요한 요소이다. 그러나 최근 산업용 IoT의 확산으로 인해 트래픽이 폭발적으로 증가함에 따라 트래픽을 효율적으로 처리해줄 할당 기법이 필요하다. 본 논문에서는 산업용 IoT 환경에서 성공적인 태스크 처리율을 높이기 위한 2단계 태스크 오프로딩 결정 기법을 제안한다. 또한, 컴퓨팅 집약적인 태스크를 셀룰러 링크를 통해 이동 엣지 컴퓨팅(Mobile Edge Computing: MEC) 서버로 오프로드 하거나 D2D(Device to Device) 링크를 통해 근처의 산업용 IoT 장치로 오프로드 할 수 있는 하이브리드 오프로딩(Hybrid-offloading) 시스템을 고려한다. 먼저 1단계는 태스크 오프로딩에 참여하는 기기들이 이기적으로 행동하여 태스크 처리율 향상에 어려움을 주는 것을 방지하기 위해 인센티브 메커니즘을 설계한다. 메커니즘 디자인 중 McAfee's 메커니즘을 사용하여 태스크를 처리해주는 기기들의 이기적인 행동을 제어하고 전체 시스템 처리율을 높일 수 있도록 한다. 그 후 2단계에서는 산업용 IoT 장치의 불규칙한 움직임을 고려하여 비정상성(Non-stationary) 환경에서 멀티 암드 밴드(Multi-Armed Bandit: MAB) 기반 태스크 오프로딩 결정 기법을 제안한다. 실험 결과로 제안된 기법이 기존의 다른 기법에 비해 전체 시스템 처리율, 통신 실패율, 후회 측면에서 더 나은 성능을 달성할 수 있음을 보인다.

키워드 : 이동 엣지 컴퓨팅, D2D(Device to Device) 오프로딩, 산업용 IoT, McAfee's 이중 경매, 멀티 암드 밴드

1. 서 론

최근 사물 인터넷의 확산에 따라 산업 분야에서 5G의 핵심

기술 중 하나인 산업용 IoT(Industrial Internet of Things: IIoT)가 활발히 연구되고 있는 추세이다. 장치 간의 대규모 데이터 통신은 산업용 IoT의 기본 기능 중 하나이며, 산업용 IoT는 산업 분야의 발전을 위해 중요한 역할을 할 것으로 예상된다. 산업용 IoT의 작업은 실시간 처리에 민감한 경우가 많다. 그러나 제한된 자원으로 인해 현재의 셀룰러 네트워크는 산업용 IoT 트래픽에 대한 대규모 연결과 낮은 지연 시간을 충족할 수 없다. 이동 엣지 컴퓨팅(Mobile Edge Computing: MEC)은

※ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터 지원사업의 연구결과로 수행되었음(IITP-2023-2018-0-01799).

† 비 회 원 : 서강대학교 컴퓨터공학과 석사

†† 중 심 회 원 : 서강대학교 컴퓨터공학부 조교수

Manuscript Received : March 2, 2023

Accepted : May 6, 2023

*Corresponding Author : Kim Sung Wook(swkim01@sogang.ac.kr)

이러한 문제를 해결하기 위한 유망한 패러다임으로 간주되고 있다. 또한, 기존 클라우드 컴퓨팅(Cloud Computing: CC)과 달리, 이동 엣지 컴퓨팅은 무선 액세스 네트워크(Radio Access Network: RAN) 내 근접한 거리에 있는 모바일 기기에 클라우드 컴퓨팅 기능을 제공하는 5G 기술로 산업용 IoT에 낮은 지연 시간을 지원할 수 있다[1,2].

하지만 산업용 IoT의 막대한 트래픽을 이동 엣지 컴퓨팅 서버가 모두 감당하기에는 한계가 있다. 이러한 한계를 극복하기 위해 태스크를 이동 엣지 컴퓨팅 서버에만 요청하는 것이 아닌 태스크 처리가 필요한 기기의 주변 기기들에 태스크 처리를 요청하는 D2D(Device to Device) 태스크 오프로딩 또한 공동으로 고려하는 것이 필요하다. 그러나 이동 엣지 컴퓨팅과 D2D 태스크 오프로딩을 공동으로 고려하는 문제는 여전히 연구 중이다[3].

이런 문제를 해결하기 위해 다양한 태스크 오프로딩 기법들이 연구되었다. 먼저, 태스크 오프로딩에 참여하는 기기들이 이기적으로 행동하여 태스크 처리율 향상에 어려움을 주는 것을 방지하기 위해 게임이론 중 하나인 계약 이론과 경매 이론을 사용한 인센티브 메커니즘 설계 연구가 진행되고 있다. 다음으로 적절한 기기에 태스크 오프로딩 요청을 위한 강화 학습의 Q 러닝 알고리즘과 멀티 암드 밴딧(Multi-Armed Bandit: MAB) 뿐만 아니라 게임 이론의 잠재력 게임과 진화 게임 기반의 태스크 오프로딩 결정 기법 연구가 있다. 그러나 효율적인 태스크 오프로딩 결정 기법을 설계하기 위해서는 인센티브 메커니즘과 오프로딩 결정 기법의 결합이 필요하다[4,5].

따라서 본 논문에서 제안하는 기법은 McAfee's 메커니즘과 멀티 암드 밴딧 기반의 이동 엣지 컴퓨팅 하이브리드 오프로딩 기법으로, 태스크 오프로딩 결정을 위해 경매 이론과 강화 학습을 사용하는 기법이다. McAfee's 메커니즘과 멀티 암드 밴딧을 사용하여 이동 엣지 컴퓨팅 환경에서 효율적인 태스크 오프로딩 기법을 제안한다. 본 논문이 기여하는 바는 다음과 같다.

1. 1단계에서 이중 경매 중 McAfee's 메커니즘을 사용하여 서비스를 제공하는 기기를 위한 인센티브 메커니즘을 설계한다. 따라서, 경매 참여자인 서비스 제공 기기와 서비스 구매 기기의 이기적인 행동을 제어하여 태스크 오프로딩 참여율을 높임으로써 전체 시스템 처리율을 높인다.
2. 2단계에서는 멀티 암드 밴딧 중 디스카운티드 톰슨 샘플링(Discounted Thomson Sampling: DTS)을 사용하여 최종 오프로딩 결정을 수행한다. 또한, 본 논문에서는 기존의 디스카운티드 톰슨 샘플링의 문제점을 해결하기 위해 기존의 기법을 수정하였다. 이를 통해 각 행동에 대한 보상 예측값의 정확도가 기존 기법에 비해 향상하여 각 서비스 요청자에게 서비스 제공 기기를 최적으로 할당한다.
3. 마지막으로, 본 논문에서 제안한 기법과 다른 태스크 오프

로딩 기법과의 비교를 통해 제안된 기법이 후회(regret)와 시스템 전체 처리율 그리고 통신 실패율 측면에 있어 더 좋은 성능을 달성할 수 있음을 입증한다.

본 논문의 구성은 다음과 같다. 2장에서는 이전 오프로딩 결정 기법들을 설명하고 3장에서는 McAfee's 메커니즘과 멀티 암드 밴딧에 대해 설명한다. 4장에서는 시스템 모델과 제안한 2단계 오프로딩 기법을 자세하게 설명한다. 5장에서는 다른 오프로딩 기법들과 성능 평가를 통해 본 논문에서 제안하는 알고리즘의 우수성을 입증한다. 6장에서는 결론으로 마무리한다.

2. 관련 연구

본 장에서는 효율적인 트래픽 관리를 위한 멀티 암드 밴딧과 이중 경매를 이용한 오프로딩 관련 연구를 설명한다.

2.1 멀티 암드 밴딧을 이용한 오프로딩 기법

멀티 암드 밴딧은 주로 추천 시스템에 많이 사용되는 기법으로 Google Analytics와 Netflix에서 추천 시스템 구축 시 사용되었다. 이러한 추천 시스템을 이용하여 네트워크 분야에서는 오프로딩 결정 기법과 채널 선택 기법 등을 위해 연구되고 있다. 컴퓨팅 집약적인 산업 작업(industrial task)을 셀룰러 링크를 통해 이동 엣지 컴퓨팅 서버로 오프로드 하거나 D2D 링크를 통해 근처의 산업용 IoT 장치로 오프로드 할 수 있는 코오프로딩(co-offloading)에 대해 조사한다. 또한, 계산 지연과 시스템 비용 최소화를 목표로 멀티 암드 밴딧 기반의 코오프로딩 결정 기법을 제안한다. 이러한 코오프로딩 기법은 네트워크 정체를 피하면서 계산 지연을 최소화한다. 하지만 불규칙한 산업용 IoT 장치의 이동과 동적 환경은 네트워크 토폴로지와 채널 상태를 변화시킨다. 이를 완전히 예측하기 어렵기 때문에 불완전한 오프로드 정보를 기반으로 공동 오프로딩 결정을 해야 한다는 단점이 존재한다[6].

무선 장치는 컴퓨팅 용량과 배터리 전력 제한 문제가 있으므로 무선 장치가 셀룰러 링크를 통해 작업을 이동 엣지 컴퓨팅 서버로 오프로드 하거나, D2D 링크를 통해 근처 유휴(idle) 장치로 태스크를 오프로드 할 수 있는 이동 엣지 컴퓨팅 시스템을 고려한다. 그러나 이러한 하이브리드 모드 태스크 오프로드 방식은 무선 장치의 이동성이 불확실하다. 따라서 이동 엣지 컴퓨팅 환경에서 PHT(Page-Hinkley Test) 감지 알고리즘과 멀티 암드 밴딧을 결합한 PhBandit이라는 오프로드 알고리즘을 제안한다. 멀티 암드 밴딧의 보상을 평균 전송 속도를 기반으로 하므로 잦은 환경 변화에서 오프로딩 결정이 부정확한 문제가 존재한다[7].

2.2 이중 경매를 이용한 오프로딩 기법

현재까지 이중 경매를 이용한 여러 오프로딩 기법들이 연구되었다. 포그 컴퓨팅(Fog Computing)에서 포그 노드의 짧

은 대기 시간과 높은 처리량을 충족하지만, 자원 제약의 문제가 있다. 따라서 포그 노드가 보장된 성능으로 가능한 많은 애플리케이션을 서비스할 수 있도록 인센티브를 제공하는 이중 경매 메커니즘을 이용한 TAFS(Truthful Auction for the Fog System)를 제안한다. TAFS는 사용자의 지연 시간 임계치를 고려해 포그 노드의 자원 제약 문제를 해결하고 이론적으로 진실성, 개인 합리성, 예산 균형 등을 충족한다는 것을 증명한다. 하지만 이중 경매 진행 시, 요청하는 자원의 양을 따로 고려하지 않고 가격을 결정하여 서로 다른 요구를 가진 포그 노드를 완전히 충족시킬 수 없다는 문제가 존재한다[8].

이동 엣지 컴퓨팅 환경에서 산업용 IoT 모바일 장치가 제출한 입찰로 엣지 서버에 오프로딩을 요청하고 엣지 서버는 요청 가격으로 컴퓨팅 서비스를 제공하는 이동 엣지 컴퓨팅과 산업용 IoT 모바일 장치의 양방향 상호 작용을 모델링한다. 이를 기반으로 이동 엣지 컴퓨팅에서 동적 가격을 가진 두 가지 이중 경매 체계인 손익분기점 기반의 이중 경매와 더 효율적인 동적 가격 책정 기반 이중 경매 방식을 제안한다. 또한, 두 알고리즘이 이론적 분석을 통해 예산 균형, 개인 이익, 시스템 효율성, 진실성을 입증한다. 하지만 이동 엣지 컴퓨팅 서버와 산업용 IoT 모바일 장치 간의 오프로딩만 고려하여 트래픽이 많은 산업용 IoT 장치의 이동 엣지 컴퓨팅 환경에서 적용이 잘 안되는 문제가 있다. 또한, 산업용 IoT 모바일 장치의 이동성을 고려하지 않아 동적 환경에도 적용이 잘 안되는 문제가 존재한다[9].

3. McAfee's 메커니즘과 멀티 암드 밴딧

본 장에서는 본 논문에서 계산 작업 처리를 담당하는 서비스 디바이스들의 참여 동기 부여를 위한 McAfee's 메커니즘을 설명하고 최종 태스크 오프로딩 결정에 사용하기 위한 멀티 암드 밴딧에 관해 설명한다.

3.1 McAfee's 메커니즘

메커니즘 디자인은 게임 이론의 한 분야로 역 게임 이론이라고 한다. 그 이유는 게임 이론은 게임이라는 특정한 상황에서 참가자들의 전략을 분석하고 이에 대한 결과를 도출해내는 이론인 것에 반해 메커니즘 디자인은 설계자가 원하는 결과를 정의하고 플레이어들이 해당 결과를 도출해내도록 유도하는 방법을 설계하는 이론이다.

메커니즘 디자인 중 McAfee's 메커니즘은 이중 경매 중 하나이다. 이중 경매란 경매 참여자 중 구매자(Buyer)만 가격을 제시하는 기존의 경매 이론과 달리, 판매자(Seller)와 구매자가 모두 가격을 제시하는 경매이다. 또한, 한 명의 중개인(broker)가 있어, 구매자는 입찰가, 판매자는 요청가를 중개인에게 제출한다. 중개인은 McAfee's 메커니즘을 이용하여 각 승자 구매자와 판매자를 결정하고 지급 가격을 결정한다. McAfee's

메커니즘은 개인 합리성(Individual Rationality)과 약한 균형 예산(Weak Balanced Budget) 그리고 진실성(Truthfulness) 3가지 특성을 가진다[10,11].

3.2 멀티 암드 밴딧

멀티 암드 밴딧은 강화 학습 기법의 하나로 카지노에서 슬롯머신 투자 최적화를 위해 만들어진 기법이다. 이 기법은 agent와 여러 개의 슬롯머신의 암(arm)으로 구성되며, 각 슬롯머신의 승리 확률(winning probability)은 다르다고 가정한다. 최종적으로 하고자 하는 것은 에이전트는 각 슬롯머신의 승리 확률을 모르기 때문에 반복을 통해 승리 확률이 가장 높은 슬롯머신의 암을 당겨 투자 이익을 최대화하고자 한다.

멀티 암드 밴딧의 핵심 문제는 탐색(exploration)과 활용(exploitation) 간의 딜레마이다. 이 둘은 어느 한쪽을 극대로 추구하면 최대의 이익을 얻을 수 없기에 이 둘을 동시에 극대로 추구할 수 없는 절충(trade-off) 관계에 있다. 이러한 절충 관계를 효과적으로 해결할 수 있는 문제가 멀티 암드 밴딧 문제이다. 이 문제를 해결하기 위한 여러 탐색 기법이 존재하는데 디스카운티드 톰슨 샘플링이 그중 하나다. 디스카운티드 톰슨 샘플링은 비정상성 시스템(Non-stationary system)에서 적용을 위해 사용되며, 이 기법의 핵심 아이디어는 과거에 관찰한 보상을 디스카운트하여 시간에 따라 변하는 모델 파라미터(θ)를 예측해 나가는 것이다. 디스카운트는 현재까지 관찰된 보상에 지나간 단위 시간만큼 $\gamma \in (0,1]$ 를 곱해서 적용한다. 즉, 현재 시각에서 m 시각 전의 보상은 현재에 $r \times \gamma^m$ 만큼 적용된다. 이 아이디어를 톰슨 샘플링에 적용하면 [4]의 Algorithm 1과 같다[12].

4. 제안하는 하이브리드 오프로딩 기법

이 장에서는, 먼저 본 논문에서 고려하는 시스템 모델을 설명한다. 그 후 본 논문에서 제안하는 기법인 2단계 McAfee's 메커니즘과 멀티 암드 밴딧 기반의 하이브리드 오프로딩 기법을 설명하고 제안하는 기법이 McAfee's 메커니즘의 3가지 특성을 만족함을 증명한다.

4.1 시스템 모델

본 논문에서 고려하는 시스템 모델은 Fig. 1에 도시된 바와 같이 D2D-MEC(Device to Device-Mobile Edge Computing) 시스템을 고려한다. 하나의 D2D 커버리지 내에는 여러 디바이스가 존재하며, 각각은 Broker Device(BD), Request Device(RD), Service Device(SD)가 될 수 있다. 먼저, BD는 이중 경매 진행 시, 경매의 중개인(Broker)역할을 하는 기기를 의미한다. RD는 태스크 처리를 요청하는 기기를 의미하며, SD는 태스크를 처리해주는 서비스 기기를 의미한다. 여기서 RD는 이동 엣지 컴퓨팅 서버나 근처 SD에 의해서만 서비스되며, 한

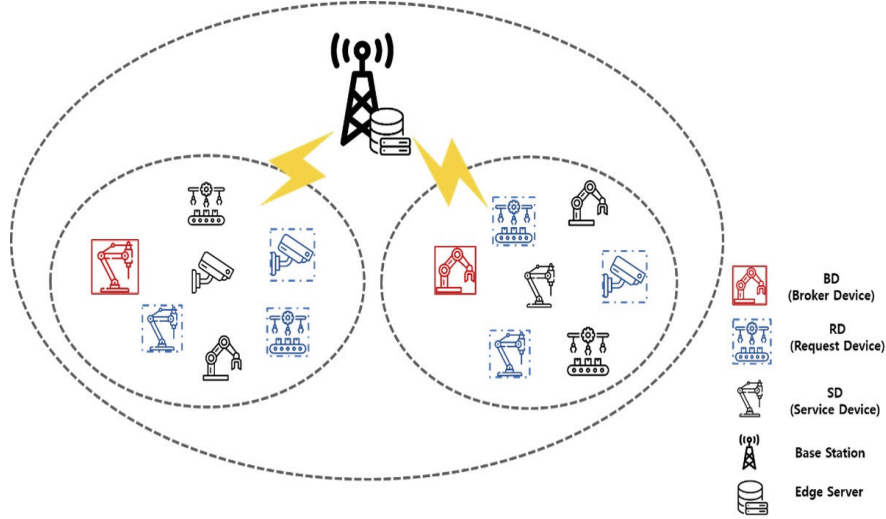


Fig. 1. System Model

번에 하나의 작업만 요청할 수 있다. 또한, 각 D2D 통신 범위 내에서 BD는 한 개이며, SD와 RD의 수는 각각 M , N 개라 가정한다[1,13].

각 시간 단계(time step) t 에서 태스크가 생성되면, 각 RD는 $(q_t^n, l_t^n, w_t^n, b_t^n)$ 을 커버리지 내의 BD에게 전송한다. 여기서 q_t^n 은 n 번째 RD가 시간 t 에서 요청하는 계산 데이터 크기(data size)이고 l_t^n 은 최대 지연시간(maximum latency)을 나타낸다. w_t^n 은 계산 작업의 작업량(workload)으로 작업 실행 완료를 위해 요청되는 CPU cycle의 수이며, b_t^n 은 n 번째 RD가 t 시간에서 이중 경매 참여를 위한 단위 입찰(unit bid)이다. 또한, 각 SD는 커버리지 내의 BD에게 단위 요청(unit ask) a_t^m 을 전송한다. 여기서 BD에게 단위 요청을 전송한 SD는 해당 시간 t 에서 유휴(idle) 상태이므로 해당 시간에 작업을 처리할 수 있다고 가정한다. BD는 이중 경매를 진행하기 전에 이동 엣지 컴퓨팅 서버로 오프로딩해야 할 RD에게 메시지를 전송한 후, 이중 경매를 통해 경매 승자를 결정하고 멀티 암드 밴딩을 이용하여 최종 오프로딩 결정 결과를 각 RD와 SD에게 전송한다. 여기서 해당 시간 t 에서 처리되지 않은 태스크는 다음 시간 $t+1$ 로 넘어가 처리되며, 태스크의 최대 지연시간(l_t^n)을 넘기는 경우 해당 태스크는 실패 처리된다.

D2D-MEC 시스템에서 RD의 계산 작업은 이동 엣지 컴퓨팅 서버나 해당 BD의 D2D 커버리지 내의 근처 SD에 의해서만 서비스되므로, RD의 D2D-MEC 오프로딩 결정을 위해 통신 가능 거리를 고려한다. 여기서 BD의 위치는 고정되어 있다고 가정한다. 또한, 각 커버리지 내의 BD와 해당 커버리지 내의 모든 RD의 최대 통신 거리는 D_{BD} 이며, RD와 BD의 거리가 D_{BD} 이상인 경우 RD의 작업은 셀룰러 링크를 통해 이동 엣지 컴퓨팅 서버로 오프로딩된다. d_t^n 은 시간 t 에서 n 번째

RD와 BD 사이의 거리를 나타내며, 거리는 유클리드 거리(Euclidean Distance)로 계산한다. RD의 오프로딩 결정 x_n 은 Equation (1)과 같다.

$$x_n = \begin{cases} 1, & d_t^n < D_{BD} \\ 0, & otherwise \end{cases} \quad (1)$$

다음은 시간 t 에서 n 번째 RD와 m 번째 SD 간의 전송률이다.

$$R_t^{n,m} = B \log_2 \left(1 + \frac{p_t^{n,m} |h_t^{n,m}|^2}{\delta^2} \right) \quad (2)$$

B 는 고정된 대역폭을 의미하며, $p_t^{n,m}$ 와 $h_t^{n,m}$ 는 시간 t 에서 n 번째 RD와 m 번째 SD 사이의 전송 전력과 평균이 0이고 분산이 σ^2 인 가우시안 분포를 따르는 채널 이득을 나타낸다. δ^2 은 가산 백색 가우스 잡음(Additive White Gaussian Noise)이다[7].

$$Energy = p_t^{n,m} \times \frac{q_t^n}{R_t^{n,m}} \quad (3)$$

Equation (3)은 멀티 암드 밴딩의 보상을 정의하기 위해 사용되는 에너지 소비량이다. 여기서 RD에 어떤 SD를 할당할지 알 수 없기 때문에 멀티 암드 밴딩 보상 정의 시 q_t^n 은 1이라고 가정한다.

실제 환경에서 f_t^m 와 채널 상태 정보를 완벽하게 알기에는 어려움이 있다. SD가 가진 계산 용량인 f_t^m 는 사적인(private) 정보이며, 채널 상태 정보는 RD와 SD의 이동성으로 인해 빠르게 변하기 때문에 실제 환경에서 위와 같은 정보를 정확하게

알기에는 어려움이 있다[14]. 따라서 본 논문에서는 이러한 문제를 해결하기 위해 RD의 정보만을 이용하여 이중 경매와 멀티 암드 밴딧 기반의 하이브리드 오프로딩 기법을 제안한다.

4.2 제안하는 기법

본 논문에서 제안하는 하이브리드 오프로딩(hybrid-off-loading) 결정 기법은 총 2단계로 구성된다. 1단계에서는 이중 경매(double auction)를 이용하여 SD들의 오프로딩 참여 동기 부여를 위한 인센티브 메커니즘을 설계하고 2단계에서는 멀티 암드 밴딧을 이용해 최종 할당을 결정한다.

1) 1단계 이중 경매 기반 인센티브 메커니즘

2단계 하이브리드 오프로딩 결정 기법의 1단계인 이중 경매 기반의 인센티브 메커니즘을 설계하기 전에 각 기기의 역할 분류가 필요하다. 각 D2D 커버리지 내에 BD는 하나가 필요하며, 이는 해당 커버리지에서 계산 자원이 가장 많은 기기가 선택된다. BD가 결정되면, 각 RD는 $(q_i^n, l_i^n, w_i^n, b_i^n)$ 정보를 BD에게 전송하고 SD는 a_i^m 정보를 BD에게 전송한다. 여기서, 시간 t 에 이중 경매에 참여하는 RD와 SD의 집합은 각각 R_t, S_t 이다. 또한, RD의 요청을 셀룰러 링크를 통해 태스크를 처리할지 D2D 링크를 통해 태스크를 처리할지 결정은 Equation (1)과 같이 요청하는 기기의 위치를 통해서 결정한다. 모든 결정이 완료되면 BD는 RD와 SD가 제공한 정보를 가지고 이중 경매를 진행한다[8].

BD는 RD와 SD에게 수신받은 정보를 가지고 RD의 b_i^n 은 내림차순으로 SD의 a_i^m 은 오름차순으로 정렬한다. 즉, $b_1^n \geq b_2^n \geq \dots \geq b_t^n$ (입찰가), $a_1^m \leq a_2^m \leq \dots \leq a_t^m$ (요청가) 이처럼 정렬한다. 그 후, $b^k \geq a^k$ 이며, $b^{k+1} < a^{k+1}$ 을 만족하는 인덱스(index) k 를 찾는다. 이 인덱스 k 는 손익분기점(breakeven index)이 되며, 이를 이용하여 경매의 승자들을 결정할 수 있다. McAfee's 메커니즘은 해당 k 를 이용하여 손익분기점 사이 적절한 가격인 $p = \frac{(b^{k+1} + a^{k+1})}{2}$ 를 계산한다. 이러한 p 가 a^k 보다 크거나 같고 b^k 보다 작거나 같은 경우 즉, $p \in [a^k, b^k]$ 인 경우 p 가 최종 가격이 되며, 1부터 k 번째 인덱스를 가진 RD와 SD가 경매의 승자가 된다. 반면, $p \notin [a^k, b^k]$ 인 경우 가격을 위와 마찬가지로 p 로 결정하면 경매 승자들은 손해가 발생할 수 있어 RD가 지불해야 하는 가격은 b^k , SD가 받아야 할 가격은 a^k 로 최종 가격이 되며, 1부터 $k-1$ 번째 인덱스를 가진 RD와 SD가 경매의 승자가 된다. 또한, 여기서 BD는 $(k-1) \times (b^k - a^k)$ 의 효용을 얻게 된다. 이와 같은 이중 경매의 결과를 이용하여 RD와 SD의 최종 할당은 2단계에서 결정된다[11].

본 논문의 이중 경매가 이상적인 메커니즘임을 증명하기 위해 3가지 특성을 만족함을 증명한다[8].

Theorem 1. 개인 합리성(Individual Rationality)

본 정리는 본 논문의 이중 경매가 모든 거래 구매자와 판매자의 개인 합리성을 만족함을 증명한다. 즉, 모든 거래 구매자와 판매자가 이중 경매에 참여함으로써 얻을 수 있는 효용은 0 이상임을 만족한다.

Proof :

내림차순 정렬된 모든 RD의 입찰과 오름차순 정렬된 모든 SD의 요청은 Equation (4)와 같은 순서로 정렬되어 있다.

$$\begin{aligned} b^1 &\geq b^2 \geq \dots \geq b^N \\ a^1 &\leq a^2 \leq \dots \leq a^M \end{aligned} \quad (4)$$

여기서 손익분기점 k 와 p 를 이용하여 경매 가격을 구하면, $p \in [a^k, b^k]$ 인 경우 각 RD가 지급해야 하는 가격과 SD가 받을 수 있는 가격은 $p = P^{RD} = P^{SD}$ 로 동일하다. 반면, $p \notin [a^k, b^k]$ 의 경우 각 RD가 지불해야 할 가격은 $b_k = P^{RD}$, 각 SD가 받아야 하는 가격은 $a_k = P^{SD}$ 이다. 따라서, 이중 경매의 RD와 SD의 승자 입찰 집합은 B_W, A_W 이며, Equation (5)와 같다.

$$\begin{aligned} B_W &= \begin{cases} b^1 \geq b^2 \geq \dots \geq b^k & \text{if } p \in [a^k, b^k] \\ b^1 \geq b^2 \geq \dots \geq b^{k-1} & \text{if } p \notin [a^k, b^k] \end{cases} \\ A_W &= \begin{cases} a^1 \leq a^2 \leq \dots \leq a^k & \text{if } p \in [a^k, b^k] \\ a^1 \leq a^2 \leq \dots \leq a^{k-1} & \text{if } p \notin [a^k, b^k] \end{cases} \end{aligned} \quad (5)$$

위와 같이 모든 B_W 에 속하는 RD는 제출한 입찰보다 작거나 같은 가격을 지불하므로 0보다 큰 효용을 얻을 수 있다.

$$b^n - P^{RD} \geq 0, \quad \forall b^n \in B_W \quad (6)$$

또한, 모든 A_W 에 속하는 SD는 제출한 요청보다 크거나 같은 가격을 받을 수 있으므로 0보다 큰 효용을 얻을 수 있다.

$$P^{SD} - a^m \geq 0, \quad \forall a^m \in A_W \quad (7)$$

따라서, 본 논문의 이중 경매는 개인 합리성을 만족한다.

Theorem 2. 약한 균형 예산(Weak Balanced Budget)

본 정리는 본 논문의 이중 경매가 약한 균형 예산을 만족함을 증명한다. 즉, 경매인(Broker)이 경매로 인해 돈을 잃지 않거나, 돈을 얻는다는 것을 만족한다.

Proof :

모든 승자 RD의 입찰 집합인 $b^n \in B_W$ 는 b^k 또는 p 즉, P^{SD} 지급이 필요하고 모든 승자 SD의 요청 집합인 $a^m \in A_W$ 는 a^k 또는 p 즉, P^{RD} 를 받는다. 또한, 2단계 멀티 암드 밴딧 기반 오프로딩 결정을 통해 각 RD와 SD는 1:1 매칭이 된다. 즉,

$|B_W| = |A_W|$ 이다. 이를 이용하여 경매인이 이중 경매를 통해 얻을 수 있는 효용을 정리하면 다음과 같다.

$$|B_W| \cdot (P^{RD} - P^{SD}) \geq 0 \tag{8}$$

이로부터 경매인의 효용은 $P^{RD} - P^{SD} \geq 0$ 이 되므로 경매인은 경매 진행 후, 돈을 잃지 않거나 얻을 수 있다. 따라서, 본 논문의 이중 경매는 약한 균형 예산을 만족한다.

Theorem 3. 진실성(Truthfulness)

본 정리는 본 논문의 모든 이중 경매 참여자는 진실 전략이 지배적 전략이며, '거짓 입찰 시 효용을 개선할 수 없다'를 만족함을 증명한다.

Proof :

먼저, 구매자(RD)의 진실 입찰은 v^n , 보고된 입찰은 b^n 이며, 해당 n 번째 RD의 보고된 입찰은 가장 큰 값을 가지지 않는다고 가정한다. n 번째 RD를 제외한 나머지 RD 들의 보고된 입찰과 판매자(SD)의 보고된 요청을 정렬하면 아래와 같다.

$$\begin{aligned} b^1 &\geq b^2 \geq \dots \geq b^{N-1} \\ a^1 &\leq a^2 \leq \dots \leq a^M \end{aligned} \tag{9}$$

Equation (9)에서 $b^k \geq a^k$ 와 $b^{k+1} < a^{k+1}$ 을 만족하는 손익분기점 k 를 찾고, $p = \frac{1}{2}(b^{k+1} + a^{k+1})$ 을 구하면 n 번째 RD의 효용은 다음과 같다.

$$U_v^n = \begin{cases} b^n - b^k & \text{if } p \notin [a^k, b^k] \text{ and } b^n \geq b^k \\ b^n - p & \text{if } p \in [a^k, b^k] \text{ and } b^n \geq p \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Equation (4.10)에서 $b^n \geq b^k$ 와 $b^n \geq p$ 는 n 번째 RD가 해당 경매의 승자임을 의미한다. RD의 진실성을 입증하기 위해, n 번째 RD가 k 와 동일한 값을 보고한다고 가정한다. 즉, b^n 은 b^k 나 b^{k+1} 이 될 수 있으며, 이는 $b^k = b^{k+1}$ 을 의미한다. $b^{k+1} < a^{k+1}$ 에 따라 p 는 $p = \frac{1}{2}(b^{k+1} + a^{k+1}) > \frac{1}{2}(b^k + b^{k+1}) = b^k = b^{k+1}$ 이 된다. p 는 $b^k < p$ 이므로 해당 경매는 $p \notin [a^k, b^k]$ 를 만족한다. $p \in [a^k, b^k]$ 인 경우 1부터 $k-1$ 번째 RD만 경매의 승자가 되므로 n 번째 RD는 경매 낙찰자가 될 수 없다. 이를 통해 $v^n \neq b^k$ 인 RD는 낙찰되지 않을 수 있으므로 거짓 입찰로 $b^n = b^k$ 로 입찰을 보고하지 않으며, $v^n = b^k$ 인 RD는 낙찰을 위해 더 높은 가격을 제시하면 효용이 음수가 되기 때문에 $b^n = b^k$ 를 보고하는 것이 지배적 전략임을 알 수 있다[12].

마지막으로 경매 참여자가 경매 가격에 영향을 줄 수 있는

유일한 방법은 경매에서 자신을 제거하는 방법이다. 즉, 경매 낙찰자가 되지 않는 것이다. 이 경우는 2가지로 나누어 확인할 수 있다. 첫 번째는 $p \notin [a^k, b^k]$ 인 경우이다. n 번째 RD는 경매 가격을 낮추기 위해 $b^n \in [a^k, b^k]$ 를 보고하여 가격을 낮출 수 있으나, $p \notin [a^k, b^k]$ 인 경우 1부터 $k-1$ 번째 RD만 경매의 승자가 된다. 따라서, n 번째 RD는 경매 낙찰자가 될 수 없다. 두 번째는 $p \in [a^k, b^k]$ 인 경우이다. n 번째 RD는 경매 가격을 낮추기 위해 $b^n < p$ 를 보고하여 가격을 낮출 수 있으며, 첫 번째 경우와 마찬가지로 경매 낙찰자에서 제외될 수 있다. k 가 손익분기점인 경우 $b^n \in [s^k, p]$ 를 만족해야 경매 가격을 낮출 수 있다. 이 경우 k 보다 더 높은 인덱스를 갖게 되어 경매 낙찰자에서 제외될 수 있다. 따라서, 본 경매에서 경매 참여자는 경매 가격에 영향을 주려는 경우 경매 낙찰자에서 제외될 수 있기 때문에 경매 참여자들은 경매에서 경매 가격에 영향을 줄 요인이 없다. SD도 위와 동일하다 [11].

본 논문에서 사용한 이중 경매인 McAfee's 메커니즘은 구매자와 판매자 간의 협의를 통해 효율적인 경매 가격을 도출할 수 있다. 또한, 경매 참여자의 진실성과 균형 예산 그리고 개인 합리성 3가지 특성을 만족시키므로 경매 참여자의 거짓 보고를 막을 수 있고 경매 참여로 인해 음수의 효용을 얻는 참여자가 존재하지 않는다. 따라서 McAfee's 메커니즘 적용을 통해 SD에 인센티브를 부여함으로써 서비스 제공 참여율을 향상시킬 수 있다.

2) 2단계 멀티 암드 밴딩 기반 오프로딩

2단계 하이브리드 오프로딩 결정 기법의 2단계인 멀티 암드 밴딩 기반 오프로딩 결정 기법은 1단계인 이중 경매 기반 인센티브 메커니즘의 이중 경매 승자 집합을 이용하여 오프로딩 작업 처리 성공률을 높일 수 있는 효율적인 오프로딩 결정 기법이다. 또한, 멀티 암드 밴딩의 탐색 기법 중 비정상성(non-stationary) 시스템을 위한 새로운 디스카운티드 톰슨 샘플링(Discounted Thomson Sampling: DTS)을 이용한다.

본 논문에서 멀티 암드 밴딩의 에이전트는 1명으로 경매 승자인 RD를 나타내며, 암(arm)은 1단계 이중 경매 기반 인센티브 메커니즘의 결과인 $S_W = \{s_1, s_2, \dots, s_K\}$ 에 속하는 SD이다. 에이전트는 SD를 선택하여 0이거나 1인 보상만 얻을 수 있다. 0의 보상은 패배(lose)를 나타내며, 1의 보상은 승리(win)를 나타낸다. 즉, 보상은 승리 확률을 의미한다. 또한, 모든 SD는 각각의 베타 분포를 가지고 있으며, 초기에 각 SD의 α, β 는 같다고 설정한다. 보상은 SD의 개인(private) 정보인 f_t^m 과 Equation (11)으로 나타내며, 보상을 0과 1 사잇값으로 표현하기 위해 정규화하여 설정한다. 에너지 소비량은 적을수록 좋으므로 분모에 계산 용량은 클수록 좋으므로 분자에

넣어 보상 식을 설정하였으며, 에이전트는 이를 알지 못한다 가정한다.

$$Reward = \frac{f_t^m}{Energy} \quad (11)$$

본 논문에서 디스카운티드 톱슨 샘플링에서 각 암의 보상을 표현하기 위해 베타 분포를 사용하였다. 그러나 기존의 디스카운티드 톱슨 샘플링은 낮은 확률을 가진 암이 초기에 높은 승리 확률을 얻게 된 후, 탐색이 덜 되면 잘 적용이 안 되는 문제가 존재한다. 따라서 본 논문에서는 탐색 초기에 다양한 암을 탐색하기 위해 기존의 디스카운티드 톱슨 샘플링의 베타 분포 샘플링 값에 소프트맥스 함수를 빼주는 방식을 추가한다. 소프트맥스 함수는 각 암이 현재까지의 반복에서 최댓값일 확률을 나타낸다. 즉, 해당 암이 현재까지 선택된 확률을 나타내는데 이를 기존의 베타 분포 샘플링 값에서 빼줌으로써 현재까지 선택이 덜 된 암이 선택될 확률을 높여주어 다양한 암 탐색이 가능해진다.

본 논문에서는 각 SD가 t 번째 반복에서 보상이 최댓값일 확률을 구하기 위해 소프트맥스 함수를 이용한다. 해당 소프트맥스 함수인 f_a 는 다음과 같다.

$$f_a = \frac{e^{\frac{\text{maximum}(a)}{T}}}{\sum_{i=1}^K e^{\frac{M(i)}{T}}} \quad (12)$$

$\text{maximum}(a)$ 는 a 번째 SD가 현재까지 반복에서 최댓값을 가진 횟수이고 $M(i)$ 는 i 번째 SD가 현재까지 반복에서 선택된 횟수를 나타낸다. 추가적으로 소프트맥스 함수는 자연 상수(Exponential Number)를 사용하여 지수적으로 급격하게 변화하는 것을 방지해주고자 T 를 사용한다. 초반에는 적절한 탐색을 위해 T 값을 크게 설정하고 반복마다 $T = T \times 0.99$ 를 사용하여 T 값을 점차 줄여준다. 따라서 Equation (12)는 a 번째 SD가 현재까지 반복에서 최대 보상을 가질 확률을 나타내며, 탐색이 덜 된 SD 위주로 더 탐색하기 위해 추가한다. 이를 위해 베타 분포에서 샘플링한 값에서 Equation (12)를 빼주어 더 적절한 탐색을 수행한다.

$$I_t = \underset{a}{\operatorname{argmax}} \theta_a(t) - f_a \quad (13)$$

Equation (13)를 이용하여 평균 보상이 높은 SD를 선택한다. 그 후, 해당 SD가 승리인 경우 $\alpha + 1$ 을 수행하고 진 경우 $\beta + 1$ 을 수행한다. 이 과정을 반복하여 각 SD의 보상에 근사해 나가며, 최종적으로 남은 시간이 적은 RD 순서대로 높은 보상을 가진 SD를 할당한다. 각 RD의 남은 시간은 각 RD의 최대 지연시간과 태스크를 요청한 시간을 이용하여 구한다[12].

BD가 2단계 멀티 암드 밴딤을 수행하고 (S_W, R_W) 결과를 얻은 후, 인덱스별로 매칭되는 SD와 RD에게 매칭 결과를 전송한다. 매칭된 RD와 SD는 D2D 링크를 통해 작업을 처리한다. 작업이 성공적으로 처리되면, n 번째 RD는 BD에게 이중 경매를 통해 결정된 단위 가격(unit cost)과 요청한 작업 크기의 곱 즉, $P^{RD} \times q_t^r$ 을 지불한다. 또한, BD는 이중 경매를 통해 결정된 단위 가격(unit cost)과 n 번째 RD의 요청과 매칭된 SD가 처리한 요청의 작업 크기의 곱 즉, $P^{SD} \times q_t^s$ 을 SD에게 지불한다. 그 후, 남은 값은 BD의 인센티브가 된다. 해당 과정에서 작업 처리를 완료하지 못한 RD는 다음 시간 단계에서 위의 과정을 반복하여 작업을 처리한다.

5. 성능 평가

본 장에서는 본 논문에서 제안한 기법의 성능 평가를 위한 환경 구성을 설명하고 기존의 기법과의 성능 비교를 통해 제안된 기법의 우수성을 보인다.

5.1 환경 구성

본 논문은 Table 1과 같이 실험을 설정한 후 제안한 오프로딩 기법의 우수성을 입증하기 위해 기존의 디스카운티드 톱슨 샘플링과 PhtBandit(Page hinkley test Multi Armed Bandit)과의 비교를 통해 분석하였다.

5.2 알고리즘 성능 분석

Fig. 2는 참여 기기의 수에 따른 시스템 처리량을 나타낸다. 참여 기기의 수가 많아질수록 제안한 기법과 비교 기법 간의 처리량의 차이가 벌어지게 된다. 이는 McAfee's 메커니즘을 사용할 경우 SD와 RD가 진실하게 요청과 입찰을 보고하기 때문에 손해를 보는 SD와 RD가 없어서 서비스 제공 참여율이

Table 1. System Parameters

System parameters	Value
Number of IIoT	[20,100]
D2D Channel Bandwidth	5 MHz
SD Computation Capacity	[0, 1] GHz
Total Time Step	5000
Learning Time Step for Multi-Armed Bandit	50
Discount factor	0.99
α_0, β_0	1
SD's ask cost range	[3, 10]
RD's bid cost range	[0, 14]

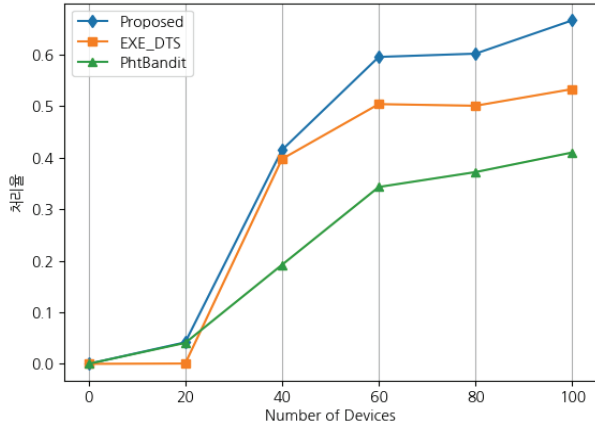


Fig. 2. Throughput by Number of Participating Devices

높아짐에 따라 전체 시스템 처리율이 높아졌다. 하지만, McAfee’s 메커니즘을 사용하지 않을 경우는 SD와 RD가 더 많은 이득을 얻기 위해 거짓 보고를 할 수 있다. 이에 따라 손해를 보는 SD와 RD가 발생하여 서비스 제공 참여율이 낮아져 전체 시스템 처리율 또한 낮아지는 결과를 보인다. 따라서 본 논문에서 제안한 오프로딩 기법은 기존의 디스카운티드 톱슨 샘플링과 PhtBandit보다 시스템 처리량에 있어 더 높은 성능을 보임을 입증했다.

Fig. 3은 참여 기기의 수에 따른 통신 실패율을 나타낸다. 실패율은 아웃티지 확률(outage probability)을 사용했다. 아웃티지 확률이란 각 수신기기가 무선 자원이 주어진 신호 대 잡음 비 임계 값인 γ_{th} 아래로 떨어질 경우 통신 중단이 발생할 확률을 나타낸다. 본 논문에서는 신호 대 잡음 비 임계 값인 γ_{th} 를 10으로 설정하였다. $\bar{\gamma}$ 는 평균 신호 대 잡음 비를 나타낸다. 각 참여 기기마다 아웃티지 확률을 구했으며, 참여 기기 수에 따라 아웃티지 확률을 평균 내어 실패율을 구한다. 아웃티지 확률은 Equation (14)와 같이 정의된다[14].

$$P_{out} = 1 - e^{-\frac{\gamma_{th}}{\bar{\gamma}}} \quad (14)$$

디스카운티드 톱슨 샘플링의 경우 매시간 단계마다 보상을 추정한다. 그러나 PhtBandit은 환경이 변하기 전 시간 단계의 보상만 고려하기 때문에 부정확한 정보로 인해 태스크 처리를 완료할 수 없는 기기에 서비스를 요청하는 경우가 생겨 디스카운티드 톱슨 샘플링에 비해 통신 실패율이 높아지는 결과를 보인다. 또한, 기존의 디스카운티드 톱슨 샘플링은 추가적인 탐색이 없어 보상을 잘못 추정하여 태스크를 완료할 수 없는 기기에 서비스를 요청하는 경우가 생겨 통신 실패율이 제안한 기법에 비해 높아지는 결과를 보인다. 따라서 통신 실패율 측면에서 제안한 알고리즘이 비교 기법들보다 더 우수한 성능을 보임을 입증했다.

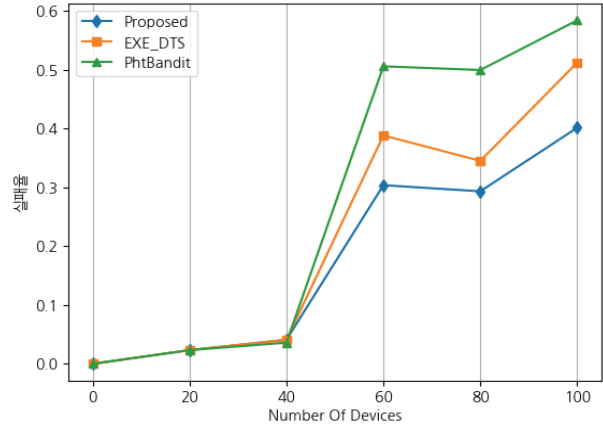


Fig. 3. Communication Failure Rate According to the Number of Participating Devices

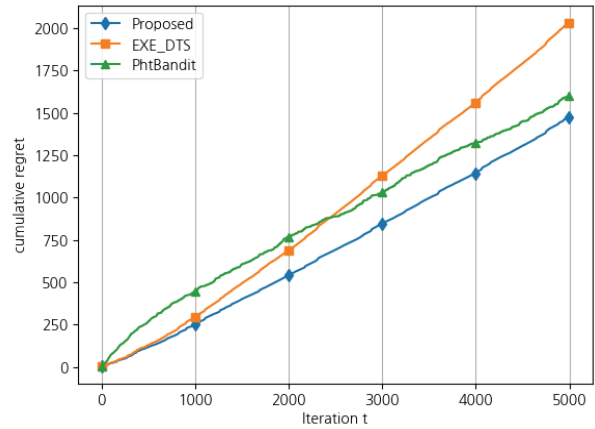


Fig. 4. Cumulative Regret According to Iteration

Fig. 4는 반복에 따른 누적 후회(Regret)를 나타낸다. 후회는 각 시간 단계에서 최적의 암 보상과 알고리즘을 이용하여 선택된 암의 보상 간 차이를 의미하며, 각 멀티 암드 밴딩 알고리즘의 예측값의 정확도를 비교하기 위해 사용된다. 따라서 누적 후회가 작은 값을 가질수록 더 높은 예측값의 정확도를 가진 알고리즘을 의미한다. 본 논문에서 암은 서비스를 제공하는 SD를 의미한다. 누적 후회는 Equation (15)와 같이 정의된다[15].

$$Regret(T) = \sum_{t=1}^T \left(\max_{1 \leq k \leq K} \theta_k - \theta_{x_t} \right) \quad (15)$$

제안한 알고리즘은 기존의 디스카운티드 톱슨 샘플링 보상 값에서 각 SD의 선택 확률을 빼줌으로써 현재까지 선택이 덜 된 SD의 선택 확률을 높여주어 다양한 SD 탐색이 가능하다. 따라서, 다양한 SD 탐색을 통해 예측값의 정확도를 높일 수 있었다. 하지만, 기존의 디스카운티드 톱슨 샘플링과 PhtBandit

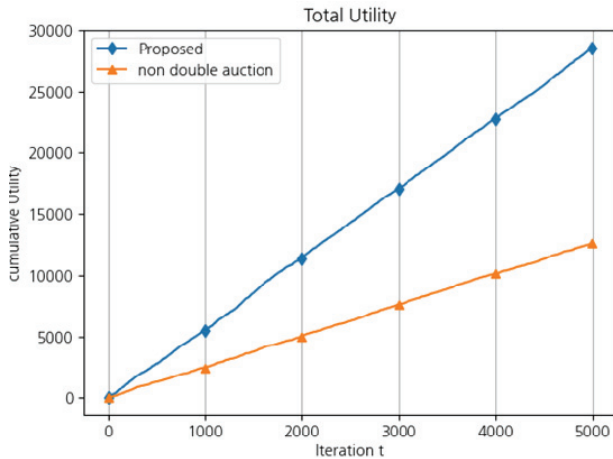


Fig. 5. Cumulative utility according to iteration

은 추가적인 탐색 없이 진행되었다. 이에 따라 낮은 보상을 가졌으나, 초기에 높은 보상을 얻은 SD의 경우 다른 SD보다 탐색이 덜 되어 보다 높은 누적 후회 값을 달성한다. 따라서 예측값의 정확도 측면에서 제안한 알고리즘이 비교 기법들보다 더 우수한 성능을 보인다.

마지막 Fig. 5는 반복에 따른 누적 효용을 나타낸다. McAfee's 메커니즘 디자인을 적용한 본 논문에서 제안한 기법과 McAfee's 메커니즘을 적용하지 않았을 경우의 효용을 비교하여 McAfee's 메커니즘을 적용할 경우 RD와 SD 그리고 BD 가 받을 수 있는 효용을 극대화할 수 있음을 보여준다. McAfee's 메커니즘을 적용할 경우 McAfee's 메커니즘의 진실 보고와 개인 합리성에 의해 RD와 SD 그리고 BD는 0 이상의 효용을 받을 수 있다. 하지만, McAfee's 메커니즘을 적용하지 않을 경우 거짓 보고를 하여 이득을 취하려 하기 때문에 RD와 SD 그리고 BD는 0보다 작은 효용을 받을 확률이 높아진다. 이를 통해 McAfee's 메커니즘을 적용할 경우 참여 기기들은 더 많은 효용을 얻을 수 있음을 입증했다.

6. 결 론

본 논문에서는 5G의 핵심 기술 중 하나인 산업용 IoT 환경에서 전체 시스템 처리 성공률 향상을 위해 이동 엣지 컴퓨팅 서버 오프로딩과 D2D 오프로딩을 고려한다. D2D 오프로딩의 경우 이기적인 장치들이 더 많은 이익을 얻기 위해 거짓 입찰을 제출하는 문제가 있을 수 있다. 또한, 실제 환경에서 장치들이 채널 상태 정보를 획득하는 데 비용이 많이 들기 때문에 채널 상태 정보를 장치들이 실제로 얻기에는 어려움이 있다[16].

위의 문제를 해결하기 위해 먼저 이중 경매를 적용하였다. 이중 경매를 적용함으로써 이기적인 장치들의 오프로딩 참여 동기 부여와 오프로딩 참여 장치들의 진실성을 보장할 수 있었으며, 구매자(RD)와 판매자(SD)의 협의를 통해 효율적인 경매 가격을 도출할 수 있었다. 또한, 실제 환경에서 장치들의

채널 상태 정보를 알기에는 어려움이 있다. 따라서 오프로딩에 참여하는 장치들은 채널 상태 정보를 모른다고 가정하며, 시간에 따라 변하는 동적인 환경에서 멀티 암드 밴딤을 사용하여 최종 오프로딩 결정을 수행하였다.

성능 평가 결과 본 논문에서 제안한 기법이 기존의 디스카운티드 톱순 샘플링과 PhtBandit 알고리즘과 비교하여 후회와 시스템 처리율 그리고 통신 실패율 측면에서 더 좋은 성능을 달성할 수 있음을 입증하였다. 추후 연구로는 각 RD가 요청하는 작업을 분할하여 병렬 처리 수행에 관한 연구와 간섭 문제를 해결하기 위한 비직교 다중 접속(Non Orthogonal Multi Access: NOMA) 기술 적용 연구가 추가적으로 필요하다.

References

- [1] N. Zhu, X. Xu, S. Han, and S. Lv, "Sleep-scheduling and joint computation-communication resource allocation in MEC networks for 5G IIoT," *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp.1-7, 2021.
- [2] X. Hou, Z. Ren, K. Yang, C. Chen, H. Zhang, and Y. Xiao, "IIoT-MEC: A novel mobile edge computing framework for 5G-enabled IIoT," *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp.1-7, 2019.
- [3] S. Koo and Y. Lim, "Optimal task offloading decision in IIoT environments using reinforcement learning," *2021 IEEE 3rd Eurasia Conference on IOT, Communication and Engineering (ECICE)*, pp.86-89, 2021.
- [4] Z. Zhou, H. Liao, B. Gu, S. Mumtaz, and J. Rodriguez, "Resource sharing and task offloading in IoT fog computing: A contract-learning approach," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, Vol.4, No.3, pp.227-240, 2020.
- [5] N. Fan, X. Wang, D. Wang, Y. Lan, and J. Hou, "A collaborative task offloading scheme in D2D-assisted fog computing networks," *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp.1-6, 2020.
- [6] J. X. Dai et al., "Task Co-offloading for D2D-assisted mobile edge computing in industrial internet of Things," in *IEEE Transactions on Industrial Informatics*, 2022.
- [7] X. Wu, T. Li, S. Yuan, Y. Lu, and Y. Wang, "A hybrid mode task offloading algorithm for edge computing," *2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp.81-85, 2021.
- [8] L. Sun, G. Xue, and R. Yu, "TAFS: A truthful auction for IoT application offloading in fog computing networks," in *IEEE Internet of Things Journal*, Vol.10, No.4, pp.3252-3263, 2022.

[9] W. Sun, J. Liu, Y. Yue, and H. Zhang, "Double auction-based resource allocation for mobile edge computing in industrial internet of things," in *IEEE Transactions on Industrial Informatics*, Vol.14, No.10, pp.4692-4701, 2018.

[10] W. Lu, S. Zhang, J. Xu, D. Yang, and L. Xu, "Truthful multi-resource transaction mechanism for P2P task off-loading based on edge computing," in *IEEE Transactions on Vehicular Technology*, Vol.70, No.6, pp.6122-6135, 2021.

[11] R. P. McAfee, "A dominant strategy double auction," *Journal of Economic Theory*, Vol.56, No.2, pp.434-450, 1992.

[12] V. Raj and S. Kalyani, "Taming non-stationary bandits: A Bayesian approach," *arXiv preprint arXiv:1707.09727*, 2017.

[13] M. Mehrabi, D. You, V. Latzko, H. Salah, M. Reisslein, and F. H. P. Fitzek, "Device-enhanced MEC: Multi-access edge computing (MEC) aided by end device computation and caching: A survey," in *IEEE Access*, Vol.7, pp.166079-166108, 2019.

[14] N. Nasaruddin, E. D. Meutia, and R. Adriman, "Outage probability and power efficiency of quantize-and-forward relay in multi-hop D2D networks," *2020 2nd International Conference on Broadband Communications, Wireless Sensors and Powering (BCWSP)*, pp.183-188, 2020.

[15] D. J., Russo et al., "A tutorial on thompson sampling," *Foundations and Trends® in Machine Learning*, Vol.11, No.1, pp.1-96, 2018.

[16] M. A. Qureshi and C. Tekin, "Fast learning for dynamic resource allocation in AI-enabled radio networks," in *IEEE Transactions on Cognitive Communications and Networking*, Vol.6, No.1, pp.95-110, 2020.



배 현 지

<https://orcid.org/0009-0004-2131-0216>

e-mail : o3od3d@sogang.ac.kr

2021년 동양미래대학교

컴퓨터소프트웨어공학과(학사)

2023년 서강대학교 컴퓨터공학(석사)

현재 삼성전자 네트워크 사업부

책임연구원

관심분야 : 게임이론을 이용한 D2D 태스크 오프로딩



김 승 욱

<https://orcid.org/0000-0003-1967-151X>

e-mail : swkim01@sogang.ac.kr

1993년 서강대학교 전자계산학과(학사)

1995년 서강대학교 전자계산학과(석사)

2003년 Syracuse University,

Computer Science(박사)

2005년 중앙대학교 컴퓨터공학부 조교수

2006년 ~ 현재 서강대학교 컴퓨터공학과 교수

관심분야 : 게임이론을 이용한 네트워크 자원관리