

무선 엣지 네트워크에서의 컴퓨팅 자원 할당 및 작업 오프로딩 기법

송희강, 구본준, 최완*

한국과학기술원, 티맥스 A&C OS연구소, *서울대학교

hghsong@kaist.ac.kr, bjgool116@gmail.com, *wanchoi@snu.ac.kr

A Study on Computing Resource Allocation and Task Offloading in Wireless Edge Network

Heekang Song, Bonjun Gu, Wan Choi*

Korea Advanced Institute of Science and Technology (KAIST), TMax A&C OS Research Institute, *Seoul National University

요약

최근 모바일 기기들의 연산이 복잡해짐에 따라 연산 작업을 근처 엣지에서 처리하는 엣지 컴퓨팅 기술이 주목받고 있다. 또한, 밀리미터파(mmWave)와 같은 차세대 통신 기술의 개발에 부응하여 무선 백홀(wireless backhaul)의 활용이 시장에서 주목받고 있다. 따라서, 본 논문에서는 무선 백홀을 사용하는 무선 엣지 네트워크에서의 엣지 컴퓨팅 자원 할당 및 사용자 작업 오프로딩 기법을 제안한다. 이러한 최적화 문제는 일반적으로 NP 난해 문제이며, 무작위 탐색을 사용하여 준 최적해를 구하는 경우라도 탐색 공간이 매우 넓어 여전히 복잡도가 높은 단점이 있다. 본 연구에서는, 적절한 제한 조건을 추가하여 전체 탐색 공간을 줄인 조건부 무작위 탐색을 제안하고, 무작위 탐색보다 더 빨리 더 좋은 준 최적해를 찾을 수 있음을 보인다.

I. 서론

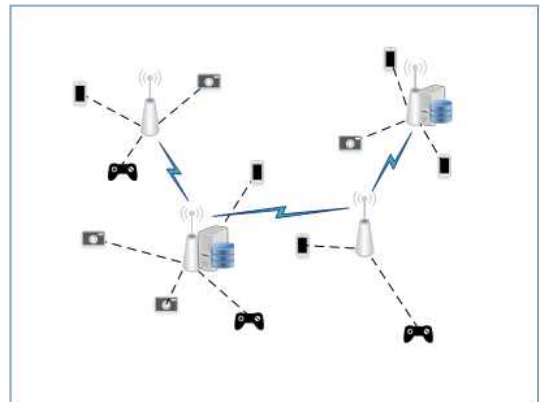
엣지 컴퓨팅은 5G 무선 네트워크 기술의 성장과 함께 디지털 혁신의 선두 분야로 부상하고 있는 기술이다. 네트워크 말단에 설치할 수 있는 서버의 연산 능력이 향상되고 더 큰 용량의 데이터를 무선으로 전송할 수 있게 되면서, 모바일 사용자들은 모바일 기기로 수집하고 생성한 데이터를 엣지 노드에 설치된 서버로 오프로드하여 서버로부터 연산 결과물을 다시 전달받는 방식의 새로운 어플리케이션을 사용할 수 있게 되었다. 대표적으로 머신 러닝, 증강 현실(AR) 및 가상 현실(VR) 등과 같이 대용량 데이터의 고속 연산을 필요로 하는 기능을 제공하는 어플리케이션들이 늘어남에 따라 모바일 기기에서 엣지 컴퓨팅 기술이 적극적으로 활용될 것으로 기대된다 [1]. 또한, 무선 지원 디바이스 개수와 데이터 트래픽이 매년 기하급수적으로 증가함에 따라, 밀리미터파(mmWave)와 같은 차세대 통신 기술의 개발에 부응하여 무선 백홀(wireless backhaul)의 활용이 시장에서 주목받고 있다 [2].

따라서, 본 논문에서는 기존 유선 백홀(wired backhaul)보다 저비용으로 광대역 전송이 가능한 무선 백홀을 사용하는 모바일 엣지 컴퓨팅을 위한 컴퓨팅 자원 및 작업 오프로딩 기법을 제안한다. 이러한 최적화 문제는 일반적으로 NP 난해 문제이며, 무작위 탐색을 사용하여 준 최적해를 구하는 경우라도 탐색 공간이 매우 넓어 여전히 복잡도가 높았지만, 본 논문에서는 적절한 제한 조건을 추가하여 전체 탐색 공간을 줄인 조건부 무작위 탐색 알고리즘을 제안하고, 시뮬레이션을 통해 제안 알고리즘이 더 빠른 시간에 좋은 성능에 도달함을 보였다.

II. 시스템 모델 및 문제 정의

본 논문에서 고려하는 무선 엣지 네트워크의 컴퓨팅 자원 할당 모델은 [그림 1]과 같이 모바일 기기들과 해당 기기들을 서비스하는 무선 AP들이 구성된다. 각 모바일 기기는 자신과 가장 가까운 무선 AP와 연결되어 있다고 가정한다. 시스템에 총 n 개의 모바일 기기들이 있고, 그들을 서비

스하는 m 개의 무선 AP가 m-means clustering algorithm을 통해 네트워크에 위치한다 [3]. 본 논문에서 해결하고자 하는 문제는 총 두 가지이다. 첫째로, 연산 작업을 처리할 수 있는 c_{total} 개의 서버 자원을 시스템 상의 m 개의 AP 중 K 개의 AP를 선택하여 선택된 AP에 얼마만큼의 서버를 설치할지에 대한 문제이다. 두번째로는 각 모바일 기기가 연산 작업을 요청할 때, 해당 작업을 오프로드 받아 처리할 서버 AP를 결정하는 문제이다. 각 문제에 대해서 컴퓨팅 자원 할당 변수 \mathbf{c} , 작업 오프로딩 변수 \mathbf{X} 를 결정한다. [그림 1]에서와 같이 AP에 연결되어 있는 모바일 기기는 작업을 연결되어 있는 AP에 무선 채널을 통해 업로드한다. 이때 연결된 AP에 서버가 설치되어 있고 서버의 컴퓨팅 자원이 충분하다면 바로 작업을 처리하지만, 그렇지 못한 AP라면 무선 백홀을 통해 근처 서버 AP에 다시 오프로드를 해주어 작업을 처리한다. 이 때, 서버 AP의 작업 컴퓨팅 시스템은 M/M/c 큐 모델에 의해 서비스해준다 [4].



[그림 1] 무선 엣지 네트워크의 컴퓨팅 자원 할당 및 작업 오프로딩 모델

본 논문에서는 컴퓨팅 자원 할당과 작업 오프로딩 최적화를 통해 평균 서비스 지연시간을 최소화하고자 한다. 각 유저의 서비스 지연 시간을 $t_i = t_{comp}(i) + t_{route}(i)$ 로 나타낼 수 있으며, $t_{comp}(i)$ 는 오프로드

한 작업을 완료하는데 걸리는 시간이고, $t_{route}(i)$ 는 연결되어 있는 AP가 서버 AP일 경우에 작업을 업로드하는데 걸리는 시간, 연결되어 있는 AP가 서버 AP가 아닌 경우, 연결되어 있는 AP에 작업을 업로드하는데 걸리는 시간과 근처 서버 AP로 오프로드하는데 걸리는 시간의 합으로 구한다. 이때 네트워크의 모든 모바일 기기의 평균 서비스 지연 시간은

$$t = \frac{1}{n} \sum_{u_i \in U} t_i \text{과 같이 나타낸다.}$$

III. 조건부 무작위 탐색 알고리즘

본 논문에서는 평균 서비스 지연 시간을 최소화하는 컴퓨팅 자원 할당 및 작업 오프로딩 기법을 찾기 위해 제한 조건을 추가한 조건부 무작위 탐색 알고리즘을 제안한다. 해당 문제는 평균 서비스 지연 시간을 최소화하는 컴퓨팅 자원 할당과 작업 오프로딩 변수를 찾는 조합 최적화 문제로, 일반적으로 풀기 어려운 NP-난해 문제이다. 따라서 모든 경우를 다 비교 해보는 완전 탐색 알고리즘은 최적해를 얻을 수 있는 탐색 방법이지만, 복잡도가 높아 해당 최적화 문제에 적용하기 적합하지 않다. 그러므로 준최적해를 얻지만 시간 복잡도를 줄이는 무작위 탐색 알고리즘을 사용하여 해를 구하는 것이 해당 문제를 푸는 방법으로 여겨진다 [5]. 본 논문에서는 기존의 무작위 탐색 알고리즘에 보다 효율적인 탐색을 위해 다음의 제한 조건을 추가하여 탐색 공간을 줄이고, 더 빠른 시간에 좋은 성능에 도달할 수 있는 조건부 무작위 탐색 알고리즘을 제안한다.

- K 개의 AP 무작위 선택
- AP에 할당하는 서버는 연결된 사용자들의 작업 요청의 합보다 크도록 할당
- 서버 AP에 연결된 각 사용자들은 작업을 오프로딩하도록 결정하고 나머지 사용자는 무작위로 오프로딩

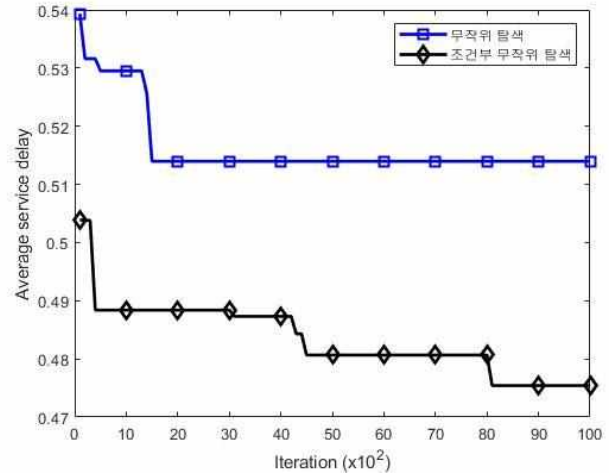
Conditioned Random Search Algorithm

```

for iteration = 1,...,L
  while 무작위로 뽑은 AP의 수가 K보다 작을 경우
    무작위로 AP 선택
    if 해당 AP가 K번째로 뽑은 AP일 경우
      해당 AP에 남아있는 컴퓨팅 자원 모두 할당
    else
      해당 AP에 연결된 사용자들의 요청의 합보다 크도록 할당
    end if
  end while (/*컴퓨팅 자원 할당*/)
  for i = 1,...,n, 모든 모바일 기기 n개에 대하여
    if 해당 사용자가 서버 AP에 연결된 사용자인 경우
      서버 AP에 해당 사용자의 작업 오프로드
    else
      무작위로 선택되고 컴퓨팅 자원이 남는 서버 AP에 해당 사용자의 작업 오프로드
    end if
  end for (/*작업 오프로딩*/)
  평균 서비스 지연 시간 t' 계산
  if 현재 iteration에서의 평균 서비스 지연 시간 t'이 이전에 계산된 최소 지연 시간 t보다 작은 경우
    최소 지연 시간 t에 현재 계산된 t' 대입
    컴퓨팅 자원 할당 및 작업 오프로딩 변수 c, X 업데이트
  end if
end for

```

제안된 알고리즘은 탐색 효율을 높여 더 빠른 시간에 더 낮은 평균 서비스 지연 시간을 달성하는 컴퓨팅 자원 할당 및 작업 오프로딩을 찾는다.



【그림 2】 탐색 공간에 따른 평균 서비스 지연 시간 비교

【그림 2】는 시뮬레이션을 통한 제안 알고리즘의 성능을 보여주고 있다. 제안된 조건부 무작위 탐색 알고리즘의 성능이 무작위 탐색보다 더 적은 탐색 공간으로 좋은 성능에 도달하는 것을 보여준다. 각 iteration마다 더 빠른 속도로 조건부 무작위 탐색이 무작위 탐색보다 낮은 평균 서비스 지연 시간을 달성하는 것을 알 수 있다.

IV. 결론

본 논문에서는 무선 엣지 네트워크에서의 컴퓨팅 자원 할당 및 작업 오프로드 기법을 연구하였다. 무선 백홀의 영향을 고려하여 평균 서비스 지연 시간을 정의하였고, 컴퓨팅 자원 할당 및 작업 오프로드를 최적화하여 평균 서비스 지연 시간을 최소화하였으며 낮은 복잡도를 가진 조건부 무작위 탐색을 제안하여 더 빠르게 최적 해를 찾을 수 있음을 보였다. 시뮬레이션 결과를 통해서 제한 조건을 추가하여 탐색 효율을 높인 제안 알고리즘이 더 빠른 시간 내에 좋은 성능에 도달함을 보였다.

ACKNOWLEDGMENT

본 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2018-0-00809, 새로운 자원을 활용한 beyond 5G 이동통신 변혁기술 개발).

참고 문헌

- [1] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.
- [2] X. Ge, H. Cheng, M. Guizani and T. Han, "5G wireless backhaul networks: challenges and research advances," in IEEE Network, vol. 28, no. 6, pp. 6-11, Nov.-Dec. 2014, doi: 10.1109/MNET.2014.6963798.
- [3] H.-H. Bock (2007) : "Clustering Methods: A History of k-Means Algorithms", Proceedings of Selected Contributions in Data Analysis and Classification, 2:161-172
- [4] L. Kleinrock, Queueing Systems, Volume 1: Theory. Hoboken, NJ, USA: pp. 101 - 103, 1975.
- [5] Zabinsky ZB (2010) Random search algorithms. Wiley Encyclopedia of Operations Research and Management Science