

분산 클라우드 환경에서 자원 활용 최대화를 위한

게임이론 기반 User Association 기법 연구

김도현[○], 홍충선^{*}
 경희대학교 컴퓨터공학과
 { doma[○], cshong^{*} }@khu.ac.kr

Game Theory based User Association Method for Resource Utilization Maximization in Distributed Cloud Environment

Do Hyeon Kim[○], Choong Seon Hong^{*}
 Department of Computer Science and Engineering, Kyung Hee University

요 약

분산 클라우드 모델은 클라우드 사용자에게 실시간, 저지연 등의 요구사항을 지니고 있는 서비스를 제공하기 위해 사용자와 가까운 곳에 에지 클라우드를 배치하는 모델로써 Access Point, Base Station 등 네트워크 액세스 인프라 상에 서버를 운용하여 서비스를 제공한다. 이러한 분산 클라우드 환경에서 사용자가 갖고 있는 요구사항(예: 컴퓨팅 자원, 서비스 품질 등)을 만족하기 위하여 어떤 에지 클라우드에 액세스해야 하는지에 대한 문제를 해결하고자, 본 논문에서는 게임이론을 활용한 기법을 제안한다. 이는 각 에지 클라우드가 각각 다른 자원량을 갖고 있으며 사용자 또한 마찬가지로 서로 다른 요구사항을 갖고 있는 상황에서, 전체 자원량의 활용도를 최대화 하는 사용자 연결 기법을 매칭 게임 알고리즘을 통하여 해결함으로써 분산 클라우드 환경을 보다 효율적으로 운용할 수 있도록 한다.

1. 서 론

스마트 디바이스 사용자가 점차 늘어나고, 서비스에 대한 요구사항이 다양해짐에 따라 클라우드 환경도 분산화 형태로 변화하고 있다. Mobile Edge Computing(MEC), Fog Computing 기술이 대표적인 분산형 클라우드 기술이라고 할 수 있다. [1]에서는 클라우드 서비스 사용자의 요구사항을 만족시키기 위하여, 사용자에게 가까이 클라우드를 배치하고 지리적으로 분산시키는 형태인 분산 클라우드 환경을 제안하였다. 에지 클라우드는 네트워크 액세스 계층에 존재하는 Access Point 또는 Base Station 상에 배치되어 사용자에게 보다 빠르고 지연이 적은 서비스를 제공하는 역할을 수행한다. 아울러 [2-3]에서는 이러한 에지 클라우드를 활용하기 위한 라우팅 모듈 및 트래픽 예측 기법 기반의 라우팅을 제안하였고 에지 클라우드로부터 서비스를 제공받을 수 있는 환경을 구축하여 분산 클라우드의 효율성에 대한 성능을 평가하였다. 이렇듯 네트워크 상에서 사용자와 가까운 곳에 배치되어 네트워크 액세스 및 서비스 제공에 활용되는 에지 클라우드를 효율적으로 운용하기 위해서는 전체 네트워크 자원을 최대로 활용하기 위한 방안을 고려하여야 한다. 본 논문에서는 이러한

문제를 해결하고자 게임이론을 활용한 에지 클라우드와 사용자 간 Association 기법을 제안한다. 이는 서로 다른 컴퓨팅 자원을 갖고 있는 에지 클라우드와 서로 다른 자원 요구사항을 갖고 있는 사용자 간의 Many-to-One 매칭 게임[4]을 통해 이루어진다. 여기서 Many-to-One 매칭 게임이 의미하는 것은 에지 클라우드가 갖고 있는 자원을 효율적으로 활용하기 위해 여러 사용자를 받아들일 수 있음을 의미 한다.

2. 관련 연구

◎ The Deferred Acceptance 알고리즘

Deferred Acceptance 알고리즘은 대학 입학 지원자와 대학 간의 매칭 또는 장기 이식 수술 대상자와 기증자 간의 매칭 같은 실생활의 문제 해결에 활용되는 알고리즘으로, 매칭이 필요한 두 집단 간의 안정적인 매칭(Stable Matching)을 위하여 사용된다. Deferred Acceptance 알고리즘의 프로세스[5]는 다음과 같다.

- ① 매칭이 필요한 두 집합 중에서 어느 한 집합의 구성원이 다른 집합의 구성원에게 제안을 하고 제안을 받은 구성원들은 각각 가장 선호하는 제안을 한 구성원을 받아들이고 나머지 구성원은 거절한다.
- ② 거절당한 구성원들은 새로운 대안들에 한해 새롭게 제안을 하고, 제안을 받은 구성원들은 추가적으로 받은 제안과 그 이전에 받은 제안 중 더 나은 하나를 다시 수락하여 받아들인다.

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학 ICT연구센터육성 지원사업의 연구결과로 수행되었음(ИTP-2018-2013-1-00717). 아울러 이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017R1A2A2A05000995). *Dr. CS Hong is the corresponding author.

③ 이러한 과정을 거쳐 더 이상 새로운 제안이 없을 시, 마지막으로 받아들인 구성원을 매칭 결과로 확정짓는다.

기존의 Deferred Acceptance 알고리즘은 Fixed Quota 즉, 정해져 있는 수용치(예: 고정된 할당량, 매칭되는 사용자에게 대한 고정된 수)를 기반으로 매칭이 이루어진다[6]. 하지만 본 논문에서는 보다 실제 시나리오에 가까운 환경을 구성하고자 에지 클라우드의 자원과 사용자의 수요량에 대하여 가변적인 값을 활용한다.

3. 제안 사항

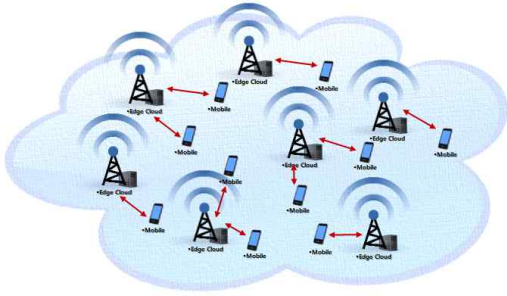


그림 1 시스템 모델

그림 1은 제안하는 알고리즘이 적용되는 시스템 모델로써, Base Station 상에 에지 서버가 배치되고 알고리즘에 의해 사용자와 에지 클라우드가 연결되는 모습을 나타낸다. 본 논문에서는 매칭 게임을 통해서 각각의 에지클라우드가 갖고 있는 컴퓨팅 자원과 각 모바일 사용자의 자원 수요량을 기반으로 전체 컴퓨팅 자원 활용을 최대화 하는 사용자 연결 기법 즉, 어떤 사용자가 어떤 에지 클라우드에 연결되어야 전체 컴퓨팅 자원을 최대한으로 활용할 수 있을지에 대한 문제를 해결하고자 한다.

$$x_{n,k}^{c_n} = \begin{cases} 1, & \text{if user } k \text{ is assigned } c_n \text{ from edge } n \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$R_{n,k}^{c_n} = \sum_{n \in N} \frac{\mu_{c_n}}{S_{c_n}} \quad (2)$$

$$U(x) = \sum_{n \in N} \sum_{k \in K} R_{n,k}^{c_n} \quad (3)$$

$$P1: \max_{x_{n,k} \in \{0,1\}} U(x) \quad (4)$$

$$\text{s.t.} \quad \sum_{k \in K} \sum_{c_n \in C_n} x_{n,k}^{c_n} \leq |C_n|, \forall n \in N \quad (5)$$

$$\sum_{c_n \in C_n} x_{n,k}^{c_n} R_{n,k}^{c_n} \geq d_k, \forall k \in K \quad (6)$$

수식 (1)은 사용자 k 에 대한 이진 변수를 나타낸 것으로 에지 클라우드 n 으로부터 컴퓨팅 자원을 할당 받으면 1, 그렇지 않으면 0으로 나타낸다. 또한 수식 (2)는 전체 에지 클라우드

자원 활용에 대한 rate로, 활용 가능한 컴퓨팅 자원 중 활용되고 있는 자원의 비율에 대한 총합을 나타낸다. 즉 모든 사용자 k 에 대하여, 아래와 같이 해결하고자 하는 문제를 수식으로 정의할 수 있다.

수식 (5)를 통해, 할당되는 컴퓨팅 자원의 양은 전체 자원 대비 작거나 같다는 제약조건을 제시하며, 수식 (6)은 각 에지 클라우드가 갖고 있는 자원과 관련하여 사용자에게 할당되는 컴퓨팅 자원은 사용자의 수요량 대비 크거나 같다는 제약조건을 나타낸다.

알고리즘 매칭 게임 알고리즘

- 1: Initially, Users and edges make their preference profiles based on their local information
- 2: At each iteration t , each user start proposing to its most preferred edge based on its profile
- 3: On receiving the proposal all edges first calculate the required computing resources
- 4: If enough computing resources are available to fulfil the computational task, it accepts the proposing users
- 5: If enough computing resources are not available, it rejects the current matched users that rank lower than the proposing users
- 6: If still sufficient computing resources are not available, the proposing users are rejected along with the previously removed users
- 7: Rejected users update their preference profile and re-propose to the second best edges in its preference profile
- 8: This iterative procedure stops if all users are either accepted by edges or there are no further edges to propose

위 알고리즘은 에지 클라우드 노드와 사용자 간의 매칭 게임 알고리즘으로써, 전체 네트워크의 컴퓨팅 리소스를 최대한으로 활용하기 위하여 어떤 사용자를 받아들여야 하는지에 대한 솔루션을 제공한다. 사용자는 모든 에지 노드에 대하여 수식 (2)를 기반으로 달성 가능한 rate를 통해 내림차순으로 정렬하여 순위를 매긴다. 아울러 에지 클라우드 노드의 경우, 자원 활용의 최대화를 이끌어 낼 수 있는 사용자 집합을 찾기 위해 수식 (3)을 기반으로 모든 사용자를 내림차순으로 정렬하여 순위를 매긴다. 그 후, 각 에지 클라우드 노드에서의 사용자 연결 수용과 거절을 통하여 매칭이 이루어지게 되며, 모든 사용자가 에지와 연결됐을 때 또는 더 이상 연결하고자 하는 에지 클라우드 노드가 없을 때 매칭 게임이 종료 된다.

이러한 과정을 거쳐 사용자와 에지 클라우드 간 매칭 게임이 이루어지며, 최종적인 결과는 에지 클라우드의 컴퓨팅 리소스가 할당된 각각의 사용자 집합으로 나타낸다.

4. 성능 평가

성능평가를 위하여, 랜덤한 사용자의 수요량 및 에지 클라우드의 자원량을 고려하였으며 500번의 시뮬레이션의 평균값을 계산하여 사용자 수용량, 전체 자원의 활용량에 대한 결과를 도출하였다. 아울러 에지 클라우드 노드 개수의 변화 변화에 따른 결과를 도출하여 알고리즘의 동작 여부 및 성능을 검증하였다.

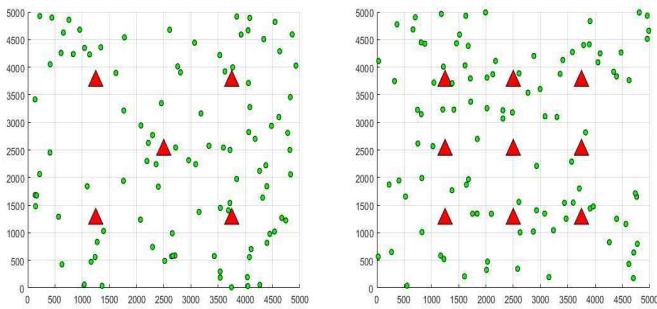


그림 2 시뮬레이션 토폴로지

그림 2는 시뮬레이션 토폴로지를 나타낸 그림으로, 다섯개와 아홉개의 에지 클라우드 노드를 배치했을 때의 성능을 분석하고자 위와 같이 환경을 구축하였다.

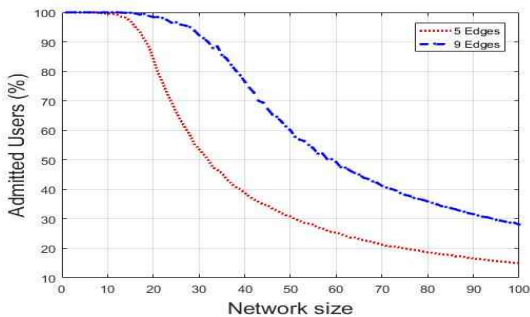


그림 3 에지 클라우드 개수에 따른 Admitted User 비율

그림 3은 에지 클라우드 노드의 수에 따른 연결된 사용자의 비율을 나타낸 그래프로써, 알고리즘에 의해 5개의 노드일 때보다 9개의 노드일 때 더 많은 사용자를 수용할 수 있는 결과를 도출하였다.

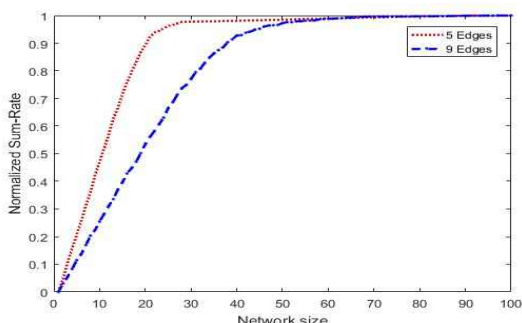


그림 4 에지 클라우드 개수에 따른 전체 자원의 Sum Rate

그림 4는 에지 클라우드 노드의 수에 따른 전체 자원의 활용 비율을 나타낸 그래프로써, 비교를 위해 자원 활용량을 정규화 하여 표현하였다. 자원 활용률의 최대값이 1이라고 할 때, 5개의 에지 클라우드가 있는 경우는 약 30명, 9개의 에지 클라우드가 있는 경우 약 60명의 사용자를 수용할 때 자원 활용률이 최대가 되는 것을 볼 수 있다. 위 두 결과를 통해 본 논문에서 제안하는 알고리즘이 잘 동작하며 자원량이 많을수록 더 많은 사용자를 수용함으로써 자원 활용률을 최대화할 수 있다는 것을 보여주고 있다.

5. 결론 및 향후 연구

본 논문에서는 에지 클라우드의 컴퓨팅 자원을 효율적으로 활용하기 위하여, 에지 클라우드 집합과 사용자 집합 간 매칭 게임 알고리즘을 제안하고 성능을 검증하였다. 매칭 게임을 통해, 전체 에지 클라우드 컴퓨팅 자원에 대한 최적의 사용자 집합을 찾음으로써 Sum Rate와 Admitted User에 대한 결과를 도출하였다.

하지만 본 논문에서 제안한 알고리즘의 경우, 에지 클라우드가 갖고 있는 컴퓨팅 리소스의 총합이 사용자 수요량보다 크거나 같다는 가정을 기반으로 한다. 이러한 이유로 향후 연구로는 코어 클라우드를 활용하여 사용자 수요량이 전체 컴퓨팅 리소스보다 큰 경우 또한 고려할 수 있는 확장된 알고리즘에 대한 연구를 진행하고자 한다.

6. 참고 문헌

- [1] YunKon Kim, Woo Sung Yun, Tien-Dung Nguyen, Eui-Nam Huh, "Study on Distributed Cloud Computing Environment with Composition model and Graph model", International Conference on Ubiquitous Information 2018 (IMCOM 2018), Jan 5-7, Malaysia, 2018.
- [2] 김도현, 김윤곤, 변준영, 손동영, Tien-Dung Nguyen, 홍충선, "분산 클라우드 환경에서 에지 클라우드 활용을 위한 서비스 인지 라우팅 모듈 구현," 2017 Korea Software Congress(2017 KSC), 12월 20일 ~ 22일, 부산, 2017
- [3] 김도현, 김윤곤, 강석원, 홍충선, "분산 클라우드 환경에서 서비스 트래픽 예측을 통한 에지 클라우드 활용 기법 연구", 2018년 한국컴퓨터종합학술대회(KCC 2018), 6월 20일 ~ 22일, 제주도, 2018
- [4] Muriel Niederle, "Matching: The Theory" online available: <https://web.stanford.edu/~niederle/MarketDesignClassIntro2.pdf>
- [5] Alvin E. Roth, "Deferred acceptance algorithms: history, theory, practice, and open questions," International Journal of Game Theory, Vol.36, No.3-4, pp.537-569, Mar 2008.
- [6] S.M. Ahsan Kazmi, Nguyen H. Tran, Walid Saad, Long Bao Le, Tai Manh Ho, Choong Seon Hong, "Optimized Resource Management in Heterogeneous Wireless Networks," IEEE Communications Letters, Vol.20, No.7, pp.1397-1400, July 2016