

D2D 네트워크 환경에서 강화학습을 이용한 사회적 관계인식 캐싱 방안에 대한 연구

임한여름^o, Kyi Thar, 홍충선^{*}

경희대학교 컴퓨터공학과

{prosummer, kyithar, cshong}@khu.ac.kr

A Research on Reinforcement Learning based Socially Aware Caching Scheme in D2D Networks

Han Yeo Reum Im^o, Kyi Thar, Choong Seon Hong^{*}

Department of Computer Science and Engineering, Kyung Hee University

요약

최근 방대한 콘텐츠를 효율적으로 서비스하기 위해 Base Station(BS)이나 User Equipment(UE)에 콘텐츠를 캐싱하는 방안이 활발히 사용되고 있다. 그러나 BS와는 달리 UE에 콘텐츠를 캐싱하기 어려운 이유는 모바일 사용자의 이기적인 특성으로 인해 다른 사용자를 위해 자신의 저장공간에 콘텐츠를 캐싱해야 할 동기 부여가 되지 않기 때문이다. 따라서 본 논문에서는 모바일 사용자 간의 사회적 관계와 D2D 링크 형성 여부를 고려하여 콘텐츠를 캐싱하도록 유도하며, Q-learning을 사용하여 전력 소비량을 낮추는 콘텐츠 배포 전략을 찾는 것을 목적으로 한다. 임의 캐싱 기법 대비 본 논문이 제시한 캐싱 기법은 훨씬 낮은 전력 소비량을 보였다.

1. 서론

최근 다양한 모바일 기기와 IoT 기기의 사용량이 급증함에 따라 넓은 대역폭에 대한 수요도 급증하게 되었다. 이에 백홀망 (backhaul link)의 혼잡도가 증가하여 모바일 트래픽이 지연되는 결과를 가져온다. 이에 콘텐츠 중심 네트워크 (Content Centric Networks)에서는 사용자의 콘텐츠 수요를 미리 파악하여 Base Stations (BSs)나 User Equipment(UE)에 캐싱하는 Proactive Caching 방안이 대두되고 있다. 일반적으로 BS에 캐싱할 땐, 해당 BS를 사용할 모바일 사용자의 수요를 예측하여 셀룰러 링크를 통해 콘텐츠를 전송하는 시나리오가 기본적이다. 반면, UE에 캐싱할 땐, 각 UE의 콘텐츠 수요 패턴, 이동 패턴 또는 UE 간의 거리 등을 고려하여 D2D 네트워크를 형성한 후 D2D 링크로 콘텐츠를 전송하는 시나리오가 일반적이다 [1].

그러나 모바일 사용자는 기본적으로 이기적이기 때문에 자신의 용량을 희생하면서 다른 사용자를 위해 콘텐츠를 캐싱하는 전략을 쉽게 받아들이기 어렵다. 따라서 본 연구는 모바일 사용자의 사회적 관계를 고려하여 사회적 관계가 높은 사용자끼리의 공동체를 형성하여 자신의 저장공간에 콘텐츠를 저장할 동기를 부여한다. 이에, D2D 네트워크를 나타낸 D2D 그래프와 사용자의 사회적 관계 그래프를 형성하여 사용자의 연결 관계를 파악한다. 이로써 콘텐츠를 각 사용자에게 최적으로 배치하여 전력 소비량을 최소화하는 것을

목적으로 하는 Q-learning 기반 캐싱 방안을 제안한다.

2. 관련 연구

■ Influence Maximization

소셜 네트워크 환경에서 널리 사용되는 Influence Maximization 기법은 사용자의 관계를 고려하여 많은 사용자에게 가장 큰 영향을 끼치는 씨드 노드를 선택한 후, 이를 활성화해 사용자 공동체에 가장 큰 영향을 끼치는 것을 목적으로 하는 해결 방안이다. 여기서 씨드 노드를 선택한 후 다른 노드들이 활성화되는 과정인 확산 과정을 다룬다. [2]에서는 공동체 내에서 하나의 씨드 노드를 선정하여 콘텐츠를 캐싱한 후, 콘텐츠의 확산 과정을 고려하여 최종적으로 영향력을 극대화하는 방안을 보여주었다.

■ Matching Game

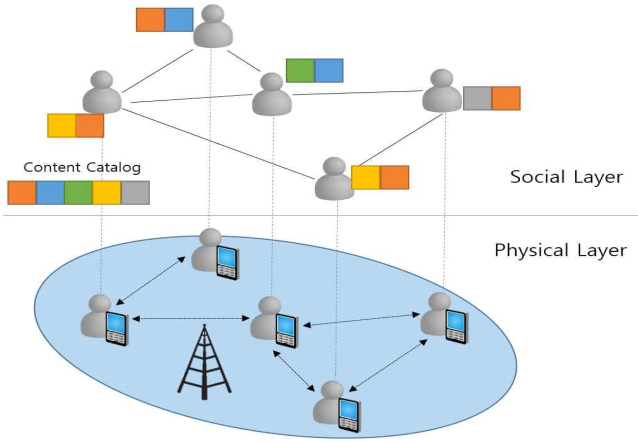
게임 이론 중 하나인 매칭게임은 이산 모델에서 노드들을 연결하는 방법론으로 일대일 매칭(결혼 문제), 다대일 매칭(대학 선택 문제), 다대다 매칭이 존재한다. 모바일 사용자를 콘텐츠를 캐싱하여 다른 사용자에게 전송할 사용자와 콘텐츠를 받기만 할 사용자로 나눌 때 사용한다. [3]에서는 콘텐츠를 캐싱할 중요한 사용자를 선택한 후 다대일 매칭으로 해당 사용자에게 콘텐츠를 받을 모바일 사용자를 조합하여 전송 지연율과 사회 복지를 증가하는 것을 목적으로 하였다.

3. 제안 사항

3.1 시스템 모델

본 논문의 시스템 모델의 예시 사례는 [그림 1]과 같다. 네트워크 내에 N 개의 기기가 존재하며 Physical

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017R1A2A2A05000995). Dr. CS Hong is the corresponding author.



[그림 1] 사회적 관계 인식 D2D 캐싱 예시 사례 Layer에서 각 사용자가 D2D 네트워크 그래프를 형성하고 있다. Social Layer에서 사용자의 공통 관심사 기반 사회적 관계 수치를 가중치로 둔 사회적 관계 그래프가 형성된다. 여기서 하나의 사회 공동체마다 콘텐츠 카탈로그는 정해져 있으며 각 콘텐츠의 크기는 L 비트로 같

다고 가정한다. 각 모바일 사용자가 콘텐츠를 요청할 때, f 번째로 인기 있는 콘텐츠를 요청할 확률은 다음의 수식 (1)과 같다.

$$p_f = \frac{f^{-\tau}}{\sum_k k^{-\tau}} \quad (1)$$

τ 는 skewness를 나타내는데, τ 가 클수록 인기 있는 콘텐츠를 요청할 확률이 더 높아진다. 본 논문의 D2D 네트워크는 inband overlaid link를 사용한다고 가정하며 따라서 셀룰러 링크와 D2D 링크의 간섭을 고려하지 않는다. 모바일 기기 m 과 n 의 채널 용량은 다음의 수식 (2)와 같다.

$$C_{m,n} = W_{m,n} \log_2 \left(1 + \frac{T_{m,n} r_{m,n}^{-\delta}}{\sigma_D^2} \right) \quad (2)$$

$W_{m,n}$ 은 모바일 기기 m, n 간의 채널 대역폭, $T_{m,n}$ 은 모바일 기기 송신 전력, 그리고 σ_D^2 은 가우시안 잡음이다. [4]와 마찬가지로 모바일 사용자 m 이 파일을 사용자 n 에게 전송할 때의 전력 소비량은 다음의 수식 (3)과 같다.

$$E_{m,n} = \frac{L}{C_{m,n}} (\beta T_{m,n} + C_D) \quad (3)$$

C_D 는 서킷 전력 소비량이며 만약 요청한 콘텐츠가 로컬에 저장되어 있다면 전력 소비량은 0이라고 가정한다. 최종적으로 본 논문에서 풀고자 하는 최적화 문제는 총 전력 소비량을 낮추는 것이므로 다음의 수식 (4)와 같다.

$$\min \left(\sum_k \sum_l^{U_N} E_{k,l} \right) \quad (4)$$

3.2 Q-learning 기반 콘텐츠 배치

강화학습의 일종인 Q-learning은 Markov Decision

[표 1] 사회적 관계 인식 D2D 네트워크 환경의 변수 및 의미

변수	의미
U_N	모바일 기기 집합
P_D	모바일 기기의 위치
r_D	모바일 기기간의 거리
r_S	사용자 간의 사회적 관계 수치
G_D	D2D 그래프 $G_D(D_N, r_D)$
G_S	사회적 관계 그래프 $G_S(D_N, r_S)$
N_f	콘텐츠 카탈로그
C_D	모바일 기기간의 채널 용량
$E_{m,n}$	사용자 m 이 n 에게 파일을 보낼 때 전력 소비량

Processes(MDPs)가 주어졌을 때 최적의 정책을 학습하기에 효과적인 학습 기법이다. 본 논문에 적합한 MDP를 정의하면 다음과 같다.

1) 상태

사회적 공동체 내의 모바일 사용자들에게 콘텐츠를 배포한 상황을 상태로 정의한다. 즉, 상태 행렬 S_{U_k, N_f} 은 각 모바일 사용자를 행으로, 콘텐츠를 열로 가진다. 예를 들어, U_2 가 파일 f_3 을 저장하고 있으면 $S_{U_2, f_3} = 1$ 이며, 그렇지 않으면 $S_{U_2, f_3} = 0$ 이다.

2) 행동

콘텐츠의 위치를 바꾸는 것을 행동으로 정의한다. 즉, 행동 행렬 A_{U_k, N_f} 은 상태 행렬과 같이 모바일 사용자, 콘텐츠를 각 행과 열로 가진다. 예를 들어, U_2 에 저장된 f_3 을 제거하고 f_5 를 저장하고 싶다면 $A_{U_2, f_3} = -1$, $A_{U_2, f_5} = 1$ 이 된다.

3) 보상

요청된 콘텐츠들을 서비스할 때의 경우에 따라 보상을 다르게 정의한다. 만약 요청된 콘텐츠가 로컬에 있는 경우는 최대 보상, one-hop 이웃에 있는 경우는 그 이웃과의 사회적 관계 수치와 거리를 고려하여 보상하며, 마지막으로 BS에서 콘텐츠를 받아야 하는 경우는 보상하지 않는다.

본 논문에서는 Q-learning으로 학습해 각 모바일 기기에 최적으로 콘텐츠를 캐싱하는 정책을 찾는 것을 목적으로 한다. 에이전트가 환경에 대해 학습한 바를 Q값에 업데이트하는 것을 수식 (5)와 같이 나타낼 수 있다.

$$Q(s,a) \leftarrow (1-\alpha)Q(s,a) + \alpha[r + \gamma \max_{a'}(Q(s',a'))] \quad (5)$$

여기서 α 는 학습률, γ 는 감소율, 그리고 s, a, r 은 각각 상태, 행동, 보상을 뜻한다. 본 논문의 Q-learning 에이전트는 ϵ 의 확률로 기존에 학습한 정책과는 다르게 임의의 행동을 하는 탐색 과정을 거친다. Algorithm 1은 Q-learning 기반 사회적 관계 인식 D2D 캐싱 기법의 수도코드를 나타낸다.

Algorithm 1. Q-learning based Socially Aware Caching Scheme

Input: D2D Graph, Social Graph
Output: Optimal Caching Policy

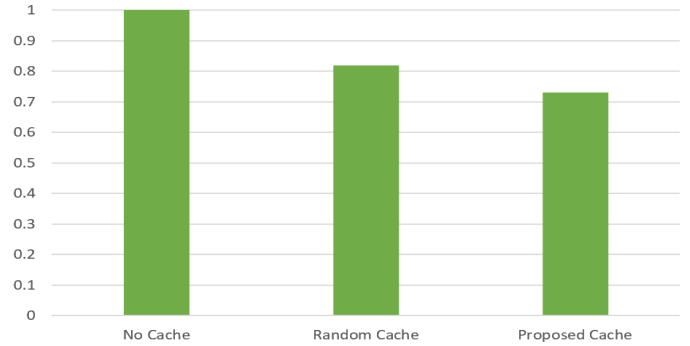
- 1: Initialize environment parameters
- 2: Set Base station with environment parameters
- 3: Define State Space and Action Space
- 4: **repeat**
- 5: **if** rand(0,1) > exploration rate ▷Exploration
- 6: choose action randomly
- 7: **else** ▷Exploitation
- 8: choose argmax action from Q table with current state
- 9: update new state, reward, Q-table
- 10: **if done then break**

4. 성능 평가

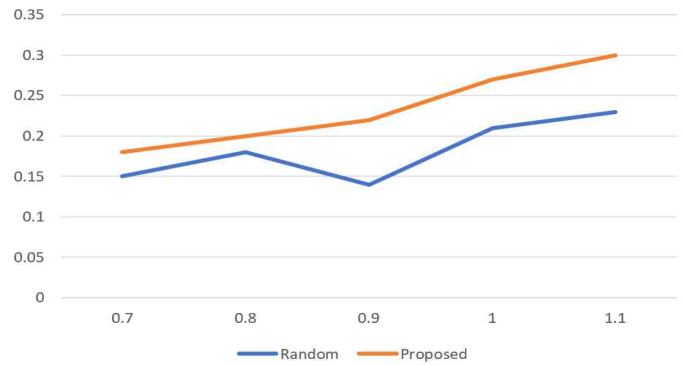
본 성능 평가는 Python 3.5를 활용하여 진행했다. 하나의 BS 범위 내의 3개의 모바일 기기가 존재하며 각 기기는 두 개의 콘텐츠를 캐시에 저장할 수 있다. 또한, 콘텐츠 카탈로그에 1000개의 콘텐츠가 있다고 가정했다. 파라미터 설정과 관련해서는 [5]을 참고했다. 제안하는 알고리즘의 성능을 평가하기 위해 모바일 기기에 콘텐츠를 임의로 저장하는 임의 캐싱 기법과 비교했다. 우측의 [그림 2]는 캐싱을 하지 않을 때의 임의 캐싱 기법과 제안한 캐싱 기법의 전력 소비량을 비교하였고 제안한 기법이 약 28%, 10% 낮은 전력 소비량을 보였다. 마지막으로, [그림 3]에서 Zipf parameter τ 에 따른 Cache Hit Ratio를 나타냈다. τ 가 커질수록 인기있는 콘텐츠의 요청 확률이 높아져 Cache Hit Ratio가 증가하는 것을 볼 수 있다. 두 그림에서 볼 수 있듯이 본 논문이 제안한 캐싱 기법이 임의 캐싱 기법보다 우월한 성능을 보임을 확인할 수 있다.

5. 결론 및 향후 연구

본 논문은 모바일 사용자 간의 사회적 관계를 고려하여 콘텐츠를 모바일 기기에 캐싱하는 하나의 방법을 제시했다. 강화학습 기반의 Q-learning을 활용하여 exploration과 exploitation을 반복, 최적의 콘텐츠 캐싱 정책을 얻을 수 있었다. 본 논문이 제시한 알고리즘이 전력 소비량의 측면에서 임의 캐싱 기법보다 월등한 성능을 보이는 것도 확인할 수 있었다. 향후 연구로 다수의 사회적 공동체를 고려하여 콘텐츠를 캐싱하는 방안과 실제 데이터를 기반으로 모바일 사용자 간의 사회적 관계 지표를 산출하는 방법을 연구함으로써 궁극적으로 Cache Hit Ratio를 높이고 지연과 전력 소비량을 줄이는 것을 목표로 연구를 진행할 계획이다.



[그림 2] 상대적 전력 소비량 비교 ($\tau = 0.9$)



[그림 3] Zipf parameter τ 에 따른 Cache Hit Ratio

6. 참고 문헌

[1] HYR. Im, CS. Hong, "A Research on Machine Learning Based Mobility Prediction Method for Content Caching", Korea Software Congress 2018, pp. 1097 - 1099, 2018

[2] M. N. Soorki, W. Saad, M. H. Manshael, and H. Saidi, "Social Community-Aware Content Placement in Wireless Device-to-Device Communication Networks," IEEE Transactions on Mobile Computing (Early Access), Aug. 2018.

[3] J. Li, M. Liu, J. Lu, F. Shu, Y. Zhang, S. Bayat, and D. N. K. Jayakody, "On Social-Aware Content Caching for D2D-Enabled Cellular Networks With Matching Theory," IEEE Internet of Things Journal, vol 6, no. 1, pp. 297 - 310, Feb. 2019.

[4] M. Chen, Y. Hao, L. Hu, K. Huang, V. K. N. Lau, "Green and Mobility-Aware Caching in 5G Networks," IEEE Transactions on Wireless Communications, vol. 16, no. 12, Dec. 2017, pp. 8347 - 8361

[5] B. Chen and C. Yang, "Energy costs for traffic offloading by cache-enabled d2d communications," IEEE Wirelss Conference and Networking Conference, Apr. 2016, pp. 1-6.