

DASH에서 Q-Learning 기반의 품질 선택 알고리즘

박지웅*, 최승식°

Q-Learning Based Quality Selection Algorithm for Dynamic Adaptive Streaming Over HTTP

Ji-Wung Park*, Seung Sik Choi°

요약

무선 디바이스와 비디오 트래픽량의 증가에 따라 무선 네트워크에서 비디오 스트리밍의 효율적인 제어가 중요한 문제로 대두되고 있다. 기존의 HTTP 적응형 스트리밍 기술들은 특정 형식에 맞춰져 있어 예측하기 어려운 네트워크 환경에 효과적으로 반응을 하지 못하며, 서로 호환이 되지 않는 문제가 있다. 이러한 문제해결을 위해 MPEG에서는 DASH(Dynamic Adaptive Streaming over HTTP)를 표준화하였고, 현재 국제 표준으로서 사용되고 있다. 본 논문에서는 DASH에서 Q-Learning을 기반으로 비디오 품질 선택 알고리즘을 제안한다. DASH에서 사용자들이 실질적으로 느끼는 체감 품질인 QoE(Quality of Experience)를 제어하기 위해 QoE 평가지표를 구성하는 항목들을 비디오 품질, 변동율과 프리즈 시간등으로 구성하고 비디오 품질 선택과 보상을 통해 사용자의 QoE 값을 최대화 할 수 있도록 설계하였다. 시뮬레이션 결과를 통해 마이크로소프트의 MSS 및 다른 논문에서 제안한 방식에 비하여 성능이 개선된 것을 확인하였으며, 기존의 방식들보다 복잡한 환경에서 개선된 결과를 보여주었다.

키워드 : 무선 네트워크, 비디오 스트리밍, 강화학습, DASH

Key Words : Wireless Network, Video Streaming, Reinforcement Learning, DASH

ABSTRACT

Efficient control of video streaming in wireless networks has emerged as an important issue with increasing wireless devices and video traffic. Traditional HTTP adaptive streaming techniques are tailored to specific formats, which do not effectively respond to unpredictable network environments, and despite their very similar techniques, there are incompatible problems. In response, MPEG proposed Dynamic Adaptive Streaming over HTTP (DASH), which would become the standard, and is now used as an international standard. To solve the problem, this paper proposes a video quality selection algorithm based on Q-Learning in DASH. Since the main goal is to meet QoE, which is the quality users feel practically, we construct Q-Value as the reward item of Q-Learning to maximize QoE. Simulation results confirm improved performance over Microsoft's MSS and other proposed methods in other papers, and show improved results in more complex environments than conventional methods, demonstrating the need for improvement in quality selection algorithms.

* 본 연구는 2018년 인천대학교 교내연구비 지원으로 수행되었습니다.

• First Author : Incheon National Univ. Department of Computer Eng., mspark1@naver.com, 학생회원

° Corresponding Author : Incheon National Univ. Department of Computer Eng., ssc Choi@inu.ac.kr, 종신회원

논문번호 : 202106-137-B-RN, Received June 21, 2021; Revised October 1, 2021; Accepted November 8, 2021

1. 서론

스마트폰과 태블릿에서 비디오 시청이 증가함에 따라, 이러한 무선 디바이스의 트래픽은 전체 인터넷 트래픽에서 비율이 증가하고 있다. 2022년까지 세계적으로 280억 개 이상의 기기와 네트워크 연결이 온라인화 되어 전 세계 인구의 60%가 인터넷을 사용하며, 이 중 동영상 트래픽 비중이 전체 트래픽의 82%를 차지할 것으로 예상된다[1]. 따라서 무선 네트워크에서 비디오 스트리밍의 효율적인 제어는 해결해야 할 중요한 문제가 될 것이다.

HAS(HTTP Adaptive Streaming)는 3세대 HTTP 기반 스트리밍이며 OTT 비디오 서비스에서 점점 더 많이 사용되고 있다. Microsoft IIS Smooth Streaming[2], Apple의 HTTP Live Streaming[3], 및 Adobe의 HTTP Dynamic Streaming[4]과 같은 몇몇 독립적인 적응형 스트리밍 기술을 구현했지만, 솔루션과 미디어 포맷의 불균형으로 인해 다양한 네트워크 상황에서 효율적인 스트리밍 서비스를 제공하기 위해 3GPP와 MPEG에서는 비디오 스트리밍의 표준이 될 DASH(Dynamic Adaptive Streaming over HTTP)를 제안했다[5].

그림 1은 DASH의 기본적인 동작방식을 나타낸다. DASH는 스트리밍 서비스의 경우 미디어 서버는 비디오 파일을 동일한 몇 초의 여러 세그먼트로 분할하고 이러한 세그먼트를 다른 여러 비트율로 서버에 저장한다. 서버는 비트율 및 URL의 세그먼트 정보를 포함하는 MPD(Media Presentation Description)를 DASH 클라이언트에 제공하고, 이러한 정보를 기반으로 DASH 클라이언트는 네트워크 상황에 따라 비디오 품질을 결정한다.

그림 2는 MPD의 구성으로서 MPD는 Period, Adaption Set, Representation, 세그먼트로 구성되어

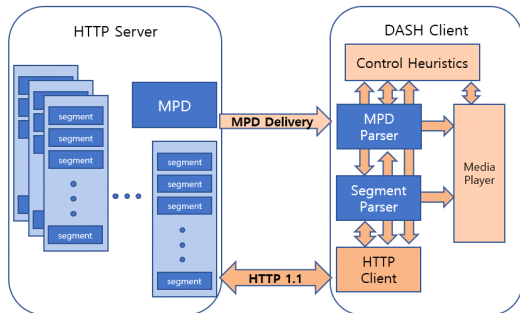


그림 1. DASH의 동작방식
Fig. 1. The basic mechanism of DASH

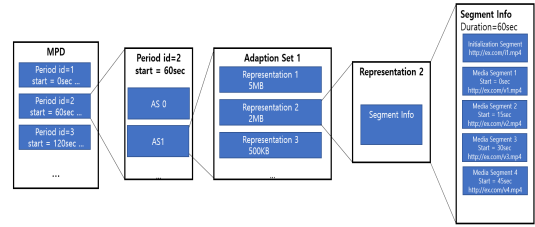


그림 2. MPD의 구성
Fig. 2. Structure of MPD

있다. Period에는 전체 미디어 프레젠테이션을 일정 시간으로 잘라 놓은 단위이다. Period는 하나 이상의 Adaption Set로 구성되는데, 같은 Period 안에서 상호 교환이 가능한 컴포넌트들의 집합을 의미한다. Representation에는 여러 품질로 압축된 미디어 컴포넌트들에 관한 정보가 있다. 각 Representation은 세그먼트들로 구성되어 있는데, 세그먼트는 HTTP에서 전송될 수 있는 가장 작은 데이터 단위이기 때문에, 그에 해당하는 URL경로가 있으며, 이 URL정보를 MPD에 기술하여 사용자는 MPD를 통해 원하는 컨텐츠를 요청할 수 있다.

DASH에 적합한 비디오 품질을 선택하려면 사용자의 QoE(Quality of Experience)를 고려하는 것이 중요하다. QoE는 서비스 사용자가 받는 서비스의 품질과 직접 관련이 있다. 따라서 QoE를 정확하게 측정하는 경우, DASH를 제어하여 QoE를 극대화하고 서비스 품질을 개선할 수 있다. 이 문제를 해결하기 위한 새로운 접근방식 중 하나인 Q-Learning을 사용하여 시스템이 학습할 수 있도록 하는 것이다. 시스템에서 하나의 state에서 하나의 action을 선택하면 action에 대한 보상이 제공된다. 이러한 action과 보상을 기반으로, 시스템이 미래에 어떤 action을 취해야 할지 알 수 있다. DASH의 경우 Q-Learning 클라이언트는 사용자의 QoE 향상을 위해 어떤 비디오 품질을 선택해야 하는지 알 수 있다.

본 논문에서는 Q-Learning 방법을 기반으로 DASH에서 사용자 QoE를 최대화 할 수 있는 방법을 제안한다. 클라이언트가 현 state에서 action을 선택하면 action에 대한 보상이 제공된다. action과 보상에 의해 계산된 Q-Value는 학습과정을 통해 계속 업데이트되며 최종 실행단계에서는 각 state에 계산된 Q-Value값을 이용하여 사용자의 QoE를 최대화하도록 비디오 품질을 선택한다. 본 논문에서는 state로 현재 비디오 품질과 세그먼트 수를 고려하고 새로운 Q-Value 계산 방법을 제시함으로써 기존 방법[13]에

비해 QoE를 향상시킬 수 있음을 결과로 제시하였다.

II. 관련 연구

DASH에서 품질 선택 알고리즘에 관한 많은 연구가 있었다. 연구는 대역폭 추정 방법, 버퍼 기반 방법, 기계 학습 방법, 수학적 모델을 사용한 방법과 같은 범주로 분류될 수 있다.

대역폭 추정 방법은 [6]과 [7]에서 사용된다. 이 접근법에서는 정확한 대역폭 추정이 핵심 요소이다. 그러나 [8]에서 HTTP에서 대역폭 추정은 본질적으로 신뢰할 수 없고 HTTP계층에서 부정확하여 비트율 변동으로 인해 낮은 품질의 비디오 스트리밍으로 이어질 수 있다고 지적하였다.

버퍼기반의 방법은 [8], [9], [10]에서 사용된다. [8]에서는 QoE에 영향을 주는 요소를 고려하여 HTTP 적응형 비디오 스트리밍을 위한 통합 사용자 QoE 모델을 제안했으며, 정적인 컨트롤러로는 최적의 서비스를 제공하기 힘들다는 것을 증명하였다. [9]에서는 버퍼 기반의 적응형 알고리즘을 제안하였다. 이 알고리즘은 초기 단계에서 가장 낮은 레벨을 선택하고 점차적으로 버퍼의 최적 레벨로 증가하여, 비디오 프리즈 시간을 줄일 수 있다. [10]은 세그먼트 수신 시간과 변화량을 기반으로 네트워크 상태를 파악하여 네트워크 상태에 따라 버퍼 임계값을 동적으로 조절하여 세그먼트 요청 시간을 스케줄링 하는 방법을 제안하였다.

수학적 모델을 활용한 기법으로는 [11], [12]로 [11]에서는 시스템의 확률적 특성을 수학적 SDP(Stochastic Dynamic Programming)문제로 특성화하여 사용자의 QoE를 극대화하는 방법을 설계한다. [11]의 방법은 좋은 결과가 나오지만 실제 시스템 동작에서 모델과 다를 경우 적응성이 떨어졌다. 이에 [12]에서는 시스템의 확률적 특성을 SDP문제로 특성화한 방법과 Q-Learning을 혼합하여 최적의 제어를 하며 Q-Learning의 적응성을 유지하였다.

마지막으로 기계학습방법을 적용한 기법으로 [11], [12], [13] 등이 있다. 3가지 방법 모두 Q-Learning 방법을 적용하여 예측하기 어려운 네트워크 환경에서 최적의 action을 학습하여 보상 값을 최대화 하여 QoE를 극대화 시키는 방법을 사용하였다.

[13]에서는 스트리밍 동작을 인식된 네트워크 상태를 기반으로 동적으로 조절할 수 있는 Q-Learning 기반 HTTP 적응형 스트리밍 클라이언트를 제안했다. 이 방식을 통해 현재의 경험적 접근 방식과 달리 클라이언트는 네트워크 환경에 해당하는 최상의 조치를

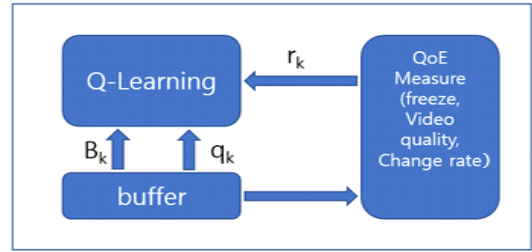


그림 3. 클라이언트 아키텍처
Fig. 3. Client Architecture

동적으로 학습할 수 있다. 이들은 HTTP 적응형 클라이언트의 동작들이 스트리밍 서비스의 QoE를 보장하기 위해서 QoE파라미터를 정의하고 이 값을 action과 보상을 이용하여 조절하기 위한 방법을 제안하였다.

[14]에서는 [11]의 모델이 시스템의 실제 동작과 일치할 경우 좋은 결과가 나오지만 실제 동작과 불일치할 경우 적응력이 부족하다 지적했으며, [13]에서 제안한 알고리즘에서 각 환경의 state는 버퍼 수준, 사용 가능한 처리량, 이전에 요청된 품질, 버퍼의 변동, oscillation의 길이와 깊이 등 6개의 변수를 사용하여 정의되어 잘못된 정책으로 이어질 수 있는 변수는 거의 없지만, 너무 많은 변수로 인해 일부 수렴 문제와 느린 학습 과정의 문제가 발생할 수 있다 지적하였다. 이에 이들은 보다 효율적으로 학습할 수 있는 방법을 얻기 위해 변수의 수를 줄이고 솔루션 성능의 핵심인 일부 함수를 정의하여 알고리즘을 개선하였다.

III. 시스템 모델

3.1 구조

Q-Learning은 기계 학습에 사용되는 강화 기법 중 하나이다. 이 기법은 에이전트가 환경에 대한 지식이 매우 제한적인 경우에 유용할 수 있다. 에이전트는 DASH 서버로부터 세그먼트를 수신하고 세그먼트를 버퍼에 저장한다. state는 버퍼의 세그먼트 수, 세그먼트 크기, 대역폭등과 같이 시스템 작업에 대해 정의된 변수로 에이전트는 state에 따라 수행할 action을 결정한다. 수행할 action은 비디오 품질의 증가, 상태유지 혹은 감소 등을 나타내는 변수로 action에 따라 다음 단계보상이 계산된다. 에이전트는 각 state에서 다음 단계보상을 계산한 후 Q-Value를 업데이트한다. 에이전트는 각 state에서 Q-Value를 사용하여 사용자 QoE를 최대화 하기 위해 수행할 action을 결정한다. 에이전트가 수행한 action은 Q-Learning에서 Q-function은 state-action 쌍의 Q-Value를 갱신하는데 사용된다.

시간 k 에서의 state와 action을 s_k 와 a_k 로 정의하고, 보상을 r 이라고 정의할 때, 시간 k 에서의 $Q(s_k, a_k)$ 는 다음과 같이 적을 수 있다.

$$Q(s_k, a_k) = (1 - \alpha)Q(s_k, a_k) + \alpha(r + \gamma \max_{a_{k+1}} Q(s_{k+1}, a_{k+1})) \quad (1)$$

이 식에서, (s, a) 는 state-action 쌍이며 $\alpha \in [0; 1]$ 와 $\gamma \in [0; 1]$ 는 각각 learning rate 와 discount factor 이다. 에이전트가 Q-Learning을 사용하여 Q-Value가 최대로 하는 action을 취하면, 서버는 변경된 품질의 세그먼트를 전송하고 에이전트가 이 action에 대한 보상을 측정한다. 그 이후에 에이전트는 이전 state의 Q-Value에 보상값을 업데이트 한다.

3.2 버퍼 모델

클라이언트의 버퍼는 서버에서 들어오는 세그먼트와 플레이어에서 나가는 세그먼트로 모델링된다. 각 세그먼트의 비디오는 여러 품질을 가지며 고정된 세그먼트의 지속 시간 동안 비디오가 재생된다. 이때 초당 재생되는 세그먼트의 재생시간을 T , 시간 k 에서 비디오 재생률을 r_k , 시간 k 에서의 버퍼의 양을 b_k 그리고 대역폭을 Bw_k 로 정의하면 b_k 는 다음과 같이 적을 수 있다.

$$b_k = b_{k-1} + \frac{Bw_k T}{r_k} - T \quad (2)$$

여기에서 만약 버퍼가 0에 달하면, 비디오 프리즈가 발생한다.

IV. DASH 기반의 Q-LEARNING

4.1 기본 접근법

DASH에서 Q-Learning을 모델링하기 위해서는 네트워크 조건의 변화에 따른 클라이언트 버퍼의 작동을 기술해야 한다. 버퍼 수준에 영향을 준 파라미터는 각 세그먼트의 비디오 품질과 대역폭이다. 그러나 대역폭은 외부 환경의 매개 변수이므로 다양한 네트워크 조건에서 측정하기가 어렵다. 특히, HTTP 트래픽의 대역폭은 TCP 트래픽 특성 때문에 예측하기가 어렵기 때문에 각 세그먼트의 품질은 내부 변수이며 세그먼트 크기에서 쉽게 측정할 수 있어야 한다. 본 논

문에서는 Q-Learning의 state 변수로 버퍼 상태와 비디오 품질을 사용한다. 시간 k 에서의 비디오 품질을 q_k 라고하면 식 (1)의 $Q(s_k, a_k)$ 는 $Q(b_k, q_k, a_k)$ 로 표현된다. 이에 따라 식 (1)을 다음과 같이 수정할 수 있다.

$$Q(b_k, q_k, a_k) = (1 - \alpha)Q(b_k, q_k, a_k) + \alpha(r + \gamma \max_{a_{k+1}} Q(b_{k+1}, q_{k+1}, a_{k+1})) \quad (3)$$

4.2 보상 함수

에이전트가 현 단계에서 한 가지 action을 취하면 다음 단계에서 보상을 측정할 수 있다. 보상으로 주어지는 결과들은 비디오 품질, 비디오 품질 변동율, 비디오 프리즈 시간으로 분류될 수 있다. 현 단계의 action이 비디오 품질의 증가인 경우 다음단계에서 비디오 품질 향상에 대한 보상이 플러스로 주어지지만 프리즈에 대한 위험이 증가하기 때문에 이에 대한 보상이 마이너스로 주어진다. 이에 따라 비디오 품질 R_q 에 대한 보상은 다음 식에 따라 계산되며,

$$R_q = \frac{q_k}{L} \quad (4)$$

L 은 비디오 품질의 개수이며, $0 < R_q < 1$ 이다. 따라서 R_q 는 비디오 품질에 따라 선형적으로 증가하며, 에이전트가 고품질 영상을 선택할 경우 더 높은 보상을 얻을 수 있다.

두 번째 보상으로 비디오 품질 변동이다. 비디오 품질 변동은 높은 품질으로의 변동으로도 불편함을 느낄 수 있기에 QoE에는 바람직하지 못한 영향을 끼친다. 따라서 비디오 품질 변동에는 마이너스의 보상이 주어진다. 비디오 품질 변동의 보상은 R_c 가 되고, 다음에 의해 계산된다.

$$R_c = \frac{|q_k - q_{k-1}|}{L} \quad (5)$$

비디오 품질 변동은 마이너스의 보상이 주어지며, 변동이 없을 시 0이 주어진다.

세 번째 보상으로 비디오 프리즈이다. 비디오 프리즈의 보상은 버퍼가 0에 가까워질 때 주어지며 비디오 프리즈는 R_f 이며 다음에 의해 계산된다.

$$R_f = \begin{cases} 1 & \text{if } b_k = 0 \\ 0 & \text{if } b_k > 0 \end{cases} \quad (6)$$

표 1. QoE와 세 가지 알고리즘들의 파라미터 비교
Table 1. QoE and parameters of three algorithms

algorithm	QoE	Q	S	F
ideal(max)	3.96	0.71	0	0
Proposed	3.92	0.73	-0.08	0
MSS	4.08	0.75	-0.04	0
Claeys	3.72	0.68	-0.05	0

총 보상 R 은 3개의 보상들의 합이 되며

$$r = C_1 R_q - C_2 R_c - (1 - C_1 - C_2) R_f \quad (7)$$

로 계산된다. 여기서 C_1 과 C_2 는 프리즈 보상에 대한 각 보상의 상대적 비율이다.

4.3 QoE 메트릭

비디오 스트리밍 서비스의 몇몇 측면을 사용자 관점에서 고려하는 것이 중요하기 때문에 이에 [14]에 정의된 QoE 메트릭을 사용한다. QoE 메트릭은 비디오 품질, 비디오 품질 변동, 비디오 프리즈의 3가지 요소로 구성된다. 비디오 품질의 평균 Q 그리고 품질 변동의 평균 S 는 다음에 의해 계산된다.

$$Q = \frac{\sum R_q}{N} \quad (8)$$

$$S = \frac{\sum R_c}{N} \quad (9)$$

이때, N 은 전체 비디오의 세그먼트 개수이다. QoE는 보상들을 사용하여 측정되며,

$$QoE = 4.85 * Q - 1.57 * S - 4.95 * F + 0.5 \quad (10)$$

로 계산되며, F 는

$$F = \frac{7}{8} * \left(\frac{\ln(F_{freq})}{6} + 1 \right) + \frac{1}{8} * \left(\frac{\min(FT_{avg}, 15)}{15} \right) \quad (11)$$

로 계산되며, F_{freq} 와 FT_{avg} 는 프리즈의 빈도와 평균 프리즈 시간을 나타낸다.

V. 실험 및 결과

5.1 파라미터 설정

모든 결과는 Python numpy 및 math 라이브러리에 시뮬레이션 되었으며, 버퍼는 4G 네트워크 사용 가능한 대역폭으로 계산되는 세그먼트 수로 모델링된다 [17]. 차량경로에서 400분 동안 측정되는 데이터를 사용하며, 시뮬레이션을 위한 트래픽 패턴은 변동성이 낮은 경우와 변동성이 높은 경우의 두 가지 조건에서 생성된다. 변동성이 낮은 경우 시뮬레이션을 위해 1Mbps의 일정한 트래픽을 사용하며, 변동성이 높은 경우 시뮬레이션을 위해, 그림 4와 같은 학습을 위한 트래픽 패턴과 그림 5와 같은 실행을 위한 트래픽 패턴을 사용한다. 대역폭 범위는 24Kbps에서 2800Kbps이며 품질의 비트율은 262, 333, 521, 789, 1030, 1242 및 2128Kbps이다. 세그먼트의 지속 시간은 2초이고 최대 버퍼 크기는 32초이며, learning rate와

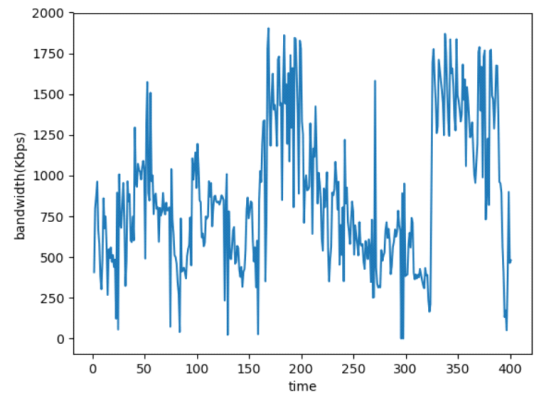


그림 4. 학습을 위한 트래픽 패턴
Fig. 4. Traffic Pattern for Exploration

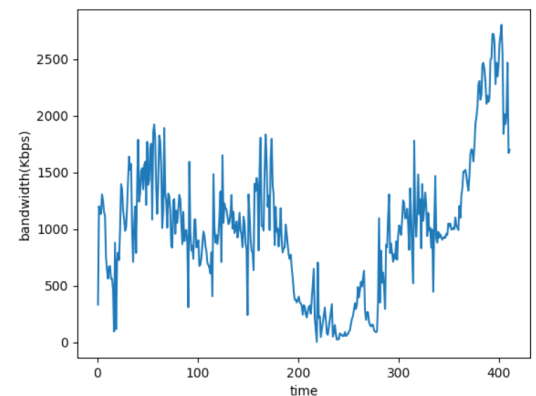


그림 5. 실행을 위한 트래픽 패턴
Fig. 5. Traffic Pattern for Exploitation

discount factor는 각각 0.1과 0.9이다. Q-Learning 탐색 단계로 200개의 에피소드를 시뮬레이션하고 마지막 에피소드를 활용 단계로 측정했다. 본 논문에서 제안한 알고리즘을 기존의 MSS 알고리즘과 Claey's^[13] 알고리즘과 비교한다.

5.2 변동성이 낮은 경우

그림 6은 제안된 알고리즘의 비디오 tx level을 보여주고 MSS와 비교한다. 그림 7에서 MSS는 1Mbps 이하의 여러 세그먼트를 수신하고 80% 버퍼 레벨을 통과하기 때문에 비디오 품질을 증가시킨다. 비디오 화질이 증가하면 20% 버퍼 수준에 도달한 다음 비디오 품질을 감소시킨다. 이 이력곡선은 그림 6과 같이 MSS의 주기를 증가시킬 수 있다. 그러나 제안된 알고리즘은 이력곡선이 나타나지 않기 때문에 트래픽 변동에 즉시 반응하여 MSS에 비해 짧은 주기를 가진다. 트래픽에 즉각적인 응답으로 인해 MSS보다 낮은 품

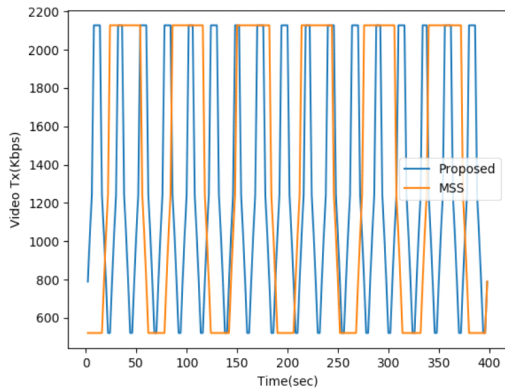


그림 6. 비디오 품질의 변화
Fig. 6. Video Quality Variation

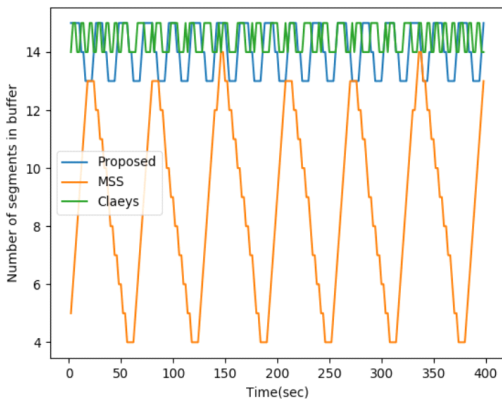


그림 7. 버퍼에서 세그먼트 수의 변화
Fig. 7. Number of Segments in buffer

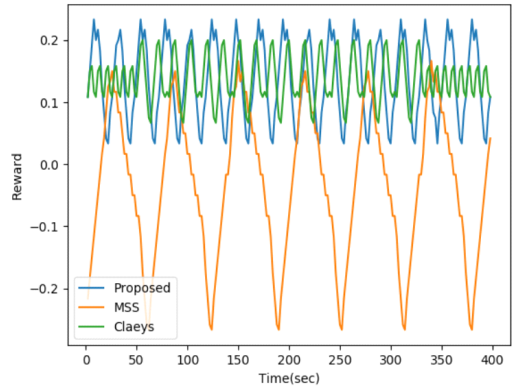


그림 8. 보상의 변화
Fig. 8. Reward Variation

질을 갖는다.

그림 7은 버퍼의 세그먼트 변화를 보여준다. MSS의 버퍼 레벨은 버퍼의 20%와 80%사이에서 오가는데, 이는 이력곡선의 특성 때문이다. 하지만 제안된 것과 Claey's의 알고리즘은 보상 값에 기반을 두어 동작한다. 이 알고리즘들은 각각의 action을 더 많은 보상 값을 얻기 위해 결정하고, MSS는 보상이 아닌 버퍼 레벨에 따라 달라진다.

그림 8에서는 식 (7)을 기반으로 세 가지 알고리즘의 각 세그먼트에서 보상의 변동을 보여준다. 제안된 알고리즘과 Claey's가 제안한 알고리즘은 보상을 기반으로 동작한다. 이 알고리즘들은 보상을 더 얻기 위해 각각의 action을 결정한다. 하지만, MSS는 보상이 아닌 버퍼 레벨에 따라 동작한다.

표 1은 낮은 변동 조건에서 마지막 에피소드의 QoE, 비디오 품질, 비디오 품질 변동 및 비디오 프리즈를 보여준다. 1Mbps 트래픽의 이상적인 QoE 값은 비디오 품질이 0.71일 때 3.96이다. 이 수치는 비디오 품질을 변경하지 않고 서비스를 제공할 때(DASH를 사용하지 않을 경우) 얻을 수 있는 최대 QoE 값을 의미한다. 비디오 품질 변경이 가능한 경우에는 MSS가 제안된 알고리즘과 Claey's의 알고리즘에 비해 QoE 값이 가장 높는데, 이는 변동성이 낮은 조건에서 MSS가 최선의 방법이라는 것을 의미한다. 또한 변동성이 낮은 조건에서는 DASH를 이용하는 MSS가 비디오 품질의 변경없이 일정한 품질의 서비스를 제공하는 경우에 비해 더 높은 QoE 값을 얻을 수 있음을 알 수 있다

5.3 변동성이 높은 경우

그림 9는 버퍼의 비디오 품질 변화를 보여준다. 그

림 10은 버퍼의 세그먼트 수를 보여준다. MSS 클라이언트에서 버퍼 범위는 중간 범위에 있으며 대역폭 변동에 적절하게 반응할 수 없다. 그러나 Q-Learning 방법은 Q-Value를 업데이트하여 이러한 다양한 네트워크 조건에 대응할 수 있다. 버퍼 상태 및 세그먼트 크기와 같은 두 가지 상태 매개 변수를 가진 제안된 방법은 더 정확하게 조치를 취할 수 있다. Claeys의 QDASH알고리즘⁸의 경우 세그먼트의 크기 대신 대역폭을 사용하였는데, 현실적으로 대역폭은 불안정하며 예측할 수 없기 때문에 이는 부정확한 Q-Value와 action의 원인이 된다. 이것은 특히 변동하는 네트워크 조건에서 프리즈 이벤트를 막을 수 없다. 그림 10은 이전 두 방식에 비하여 버퍼가 비어있는 상태를 피할 수 있음을 보여준다.

그림 11은 각 세그먼트에서 보상의 변동을 보여준다. MSS는 변동 조건이 적을 때와 마찬가지로 버퍼

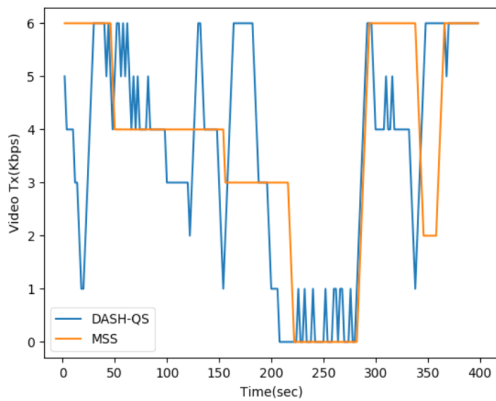


그림 9. 비디오 품질의 변화
Fig. 9. Video Quality Variation

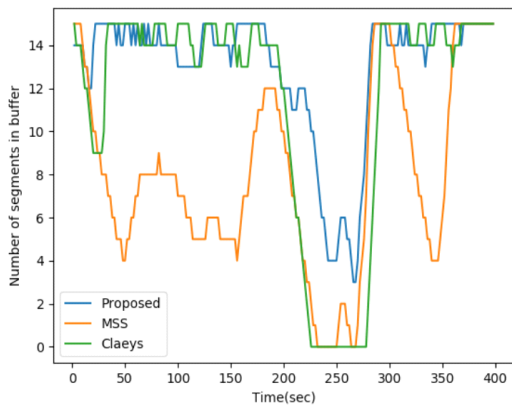


그림 10. 버퍼에서 세그먼트 수의 변화
Fig. 10. Number of Segments in buffer

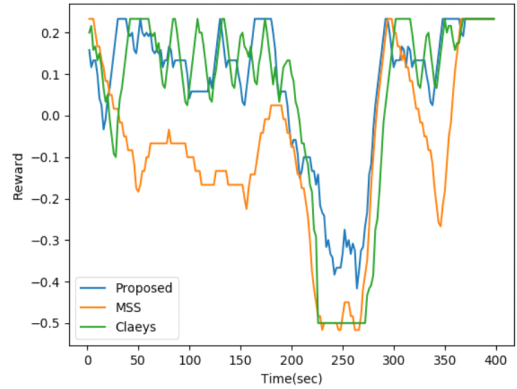


그림 11. 보상의 변화
Fig. 11. Reward Variation

표 2. QoE와 세 가지 알고리즘들의 파라미터 비교
Table 2. QoE and parameters of three algorithms

algorithm	QoE	Q	S	F
Proposed	3.32	0.63	-0.14	0
MSS	2.40	0.62	-0.07	-0.20
Claeys	1.97	0.68	-0.14	-0.32

상태에 따라서만 동작한다. 버퍼 레벨이 높아지면 비디오 품질이 높아지며, 버퍼가 비어 있다면 비디오 품질 또한 낮아진다. MSS의 보상들은 세그먼트와 대역폭이 아닌 버퍼의 상태에만 영향을 받는 것을 보여준다. 그러나 Q-learning 방법들은 다른 두 가지 변수에 의해 영향을 받는 것을 보여준다. 그림 11은 Q-Learning 기반의 방법들이 버퍼 상태뿐만 아니라 세그먼트 크기 및 대역폭에 따라 달라진다는 것을 보여준다.

표 2는 마지막 에피소드에서 세 가지 방법의 QoE, 비디오 품질, 비디오 품질 변동 및 비디오 프리즈에 대해 보여준다. 제안된 알고리즘의 QoE는 다른 알고리즘에 비해 F의 장점에 의해 높은 QoE를 갖는다.

VI. 결론

본 논문에서는 효과적인 품질 선택을 하기 위해 기존의 방법과 다른 변수를 사용하여 다른 보상 함수를 제안하였다. 제안된 방법이 실제 네트워크와 유사하게 모델링된 환경에서 평가지표가 우수하게 나온 것을 확인하였다. 또한 변동성이 낮은 경우와 변동성이 높은 경우를 각각 실험을 하여 낮은 변동 조건에서는 MSS의 성능이 더 우수하게 나왔지만 높은 변동 조건

에서는 제안한 방법의 성능이 기존방식에 비해 높은 QoE 성능을 얻을 수 있음을 알 수 있었다. 제안한 방법에서는 사용자의 QoE에 가장 치명적으로 작용하는 비디오 프리즈가 변동성이 높은 경우와 변동성이 낮은 경우 모두에서 나타나지 않았기 때문에 높은 점수를 얻을 수 있었다는 것을 실험결과를 통해 확인할 수 있었다.

References

- [1] Cisco, *Cisco Visual Networking Index: Forecast and Trends, 2017-2022*, Cisco, White Paper, 2019.
- [2] A. Zambelli, *Smooth streaming technical overview*, Microsoft Corp., Redmond, WA, USA, Tech. Rep. Retrieved May 4, 2021.
- [3] *HTTP Live Streaming overview*, Apple Inc. Cupertino, CA, USA, Retrieved May 4, 2021.
- [4] *HTTP Dynamic Streaming*, Adobe System Inc. San Jose, CA, USA, Retrieved May 4, 2021.
- [5] *Information technology—Dynamic adaptive streaming over HTTP (DASH)*, Retrieved May 5, 2021.
- [6] L. De Cicco and S. Mascolo, "An adaptive video streaming control system: Modeling, validation, and performance evaluation," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 526-539, Apr. 2014.
- [7] J. Jiang, V. Sekar, and H. Zhang, "Improving fairness, efficiency, and stability in HTTP-based adaptive video streaming with FESTIVE," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, pp. 326-340, Jan. 2014.
- [8] W. Hung, Y. Zhou, X. Xie, and D. Wu, "Buffer state enough: Simplifying the Design of QoE-Aware HTTP Adaptive Video Streaming," *IEEE Trans. Broadcasting*, vol. 64, no. 2, pp. 1-12, Dec. 2017.
- [9] H. Le, D. Nguyen, N. Ngoc, A. Pham, and T. C. Thang, "Buffer-based bitrate adaptation for adaptive HTTP streaming," in *Proc. IEEE Int. Conf. Advanced Technol. Commun.*, pp. 33-38, Ho Chi Minh, Vietnam, Oct. 2013.
- [10] D. Yun and K. Chung, "Segment scheduling scheme to support seamless DASH-based live streaming service," *KIISE Trans. Computing Practices*, vol. 22, no. 7, pp. 310-314, Jul. 2016.
- [11] S. Garcia, J. Cabrera, and N. Garcia, "Quality-optimization algorithm based on stochastic dynamic programming for mpeg dash video streaming," in *2014 IEEE ICCE*, pp. 574-575, Las Vegas, USA, Jan. 2014.
- [12] S. Garcia, J. Cabrera, and N. Garcia, "Evaluation of Q-Learning approach for HTTP adaptive streaming," *IEEE Int. Conf. Consumer Electron.*, pp. 293-294, Las Vegas, USA, Jan. 2016.
- [13] M. Claeys, S. Latre, J. Famaey, and F. De Turck, "Design and evaluation of a self learning HTTP adaptive video streaming client," *IEEE Commun. Lett.*, vol. 18, no. 4, Apr. 2014.
- [14] V. Martin, J. Cabrera, and N. Garcia, "Q-Learning based control algorithm for HTTP streaming," *IEEE Conf. VCIP*, pp. 1-4, Singapore, Dec. 2015.
- [15] T.-Y. Huang, R. Johari, N. McKeown, M. Trunnell, and M. Watson, "A buffer-based approach to rate adaptation: Evidence from a large video streaming service," *ACM Conf. SIGCOMM*, pp. 187-198, Chicago, USA, Aug. 2014.
- [16] R. Mok, E. Chan, and R. Chang, "Measuring the quality of experience of HTTP video streaming," in *IFIP/IEEE Int. Symp. Integr. Netw. Manag.(IM)*, pp. 485-492, Dublin, Ireland, May 2011.
- [17] H. Riiser, P. Vigmostad, C. Griwodz, and P. Halvorsen, "Commute path bandwidth traces from 3G networks : Analysis and applications," *ACM Conf. Multimedia Syst. (MMSys)*, pp. 114-118, Oslo, Norway, Feb. 2013.

박 지 웅 (Ji-Wung Park)



2018년 2월 : 인천대학교 컴퓨터공학부 졸업
2018년 3월~현재 : 인천대학교 컴퓨터공학부 석사과정
<관심분야> 영상스트리밍 프로토콜, 사물인터넷

최 승 식 (Seung Sik Choi)



1988년 : 연세대학교 전자공학과 졸업 (학사)
1990년 : KAIST 대학원 전기 및 전자공학과 졸업(석사)
2002년 : KAIST 대학원 전기 및 전자공학과 졸업(박사)
1990년~2004년 : KT 연구센터 선임연구원

2004년~현재 : 인천대학교 컴퓨터공학과 교수
<관심분야> 사물인터넷, 무선 MAC, 인터넷 프로토콜