**ABSTRACT**

Douyin sử dụng

* chiến lược Next-One
* dựa trên HTTP progressive download, sliding window, stop-and-wait protocol
* hoạt động kém khi kết nối mạng chậm và người dùng lướt nhanh

Liveclip

* sơ đồ phát trực tuyến thích ứng thông minh cho các video dạng ngắn trên thiết bị di động
* sử dụng deep reinforcement learning (DRL)

**INTRO**

* thực hiện một nghiên cứu đo lường chuyên sâu trên Douyin: thấy đc chiến lượng next-one kém
* áp dụng deep reinforcement learning, xem xét nhiều video trong sliding window

**A MEASUREMENT STUDY OF DOUYIN**

* các video sẽ được hiển thị toàn màn hình
* người dùng không kiểm soát các video sẽ phát, chỉ có thể chuyển đổi chế độ xem video bằng cách cuộn màn hình
* Hình 1 cho thấy tiến trình phân phối của các video liên tiếp không có người dùng
* Next-One strategy: trình phát luôn tải xuống video hiện tại, khi hoàn tất, bắt đầu tải xuống tệp video tiếp theo, Kết nối trước đó bị đóng ngay lập tức khi có sự kiện người dùng scrolling
* Hình 2 hiển thị cách chiếm dụng bộ đệm thay đổi
* sử dụng Appium: mô phỏng hành vi cuộn của người dùng thực

**DESIGN OF LIVECLIP ALGORITHM**

**Markov Decision Process**

* **T**: tổng thời gian 1 phiên xem
* bt : tốc độ download
* ct : người dùng tương tác
* dt: player state
* Tslot : khe thời gian
* Player: đưa ra quyết định, cần chọn video nào trong danh sách phát sẽ được tải xuống tại mỗi khoảng thời gian Tslot. Mỗi time step, chọn hành động **a,** chuyển sang trạng thái mới s’ từ trạng thái s cho trước xác suất chuyển đổi Pa(s,s’) và nhận được phần thưởng tương ứng Ra(s,s’)

**State S**

* {s1,s2,…..,sT} không gian mẫu state
* st = (bt,ct, dt ) : trạng thái hiện tại

1. **Player**

* Tslot và video segment : 1 second unit
* vt: video đang đc xem ở thời điểm t
* Bvt : buffered content
* Qvt : play progress
* Mvt : video length
* Player chỉ phát video khi segment hiện tại đã đc tải xong

**Text, letter

Description automatically generated**

* giả sử một video được mã hóa với tốc độ bit không đổi để mỗi phân đoạn có cùng độ dài có cùng kích thước.
* trong mỗi 1 time slot, player chỉ có thể yêu cầu một video trong sliding window

 tập video lấy về (fetched)

* Khi người dùng cuộn màn hình, nội dung video được vuốt sang bên sẽ bị loại bỏ ngay lập tức.
* Edt: tập bị loại bỏ

Text

Description automatically generated

* Nội dung đệm của các video đã hoán đổi sẽ bị xóa ngay lập tức khỏi bộ đệm để dành dung lượng cho các video trong tương lai. Với khoảng time slot là 1 giây, việc phát video sẽ dừng khi segment 1s tiếp theo chưa được lưu hoàn toàn vào bộ đệm

1. **Download speed and user interactions.**

* xây dựng lớp CNN 1 chiều tìm hiểu mô hình tốc độ tải xuống, lấy tốc độ down từ **k** giây trc làm INPUT
* xây dựng tiếp lớp CNN 1 chiều khác, tìm hiểu mô hình hành vi user, INPUT: thời gian user xem **k** video trước

1. **State space:** các thành phần

* Download speed measurement of past k seconds.
* User staying time of past k videos.
* Play progress.
* Current staying time.
* Replay round.
* Bitrate of videos in the sliding window.
* Length of videos in the sliding window.
* Buffered content of videos in the sliding window.
* Time spent downloading videos in the sliding window.
* Videos completed in the sliding window.

**Actions**

* A= {a1,a2,.., aT} không gian mẫu action
* **at** = (Eft, Edt)
* Without loss of generality, swapped videos will be evicted from the buffer immediately. (???)
* Mỗi time slot, player cần quyết định video nào trong sliding window nên được down, cũng có thể quyết định dừng tải xuống bất kỳ lúc nào để tránh lãng phí chi phí.

**Text, letter

Description automatically generated**

**State Transitions.**

* Xác suất chuyển đổi trạng thái từ st sang st+1 với action thực hiện là at

**Text, letter

Description automatically generated**

**: là xác suất do mô hình player đưa ra.**

**Reward: phụ thuộc vào (QoE, Cost)**

* **QoE:** được định nghĩa là (indicator) chỉ thị khi sự kiện rebuffering et ∈ {0, 1}, cho biết liệu video hiện tại vt có dừng phát lại do bộ nhớ cache bỏ lỡ tại thời điểm t hay không

1. rebuffering ảnh hưởng rất nhiều đến trải nghiệm người dùng

* **player** chỉ phát video khi buffer video hiện tại Bvt đi trc 1 segment so vs Qvt hoặc chạm tới hết độ dài video
* **bỏ qua chất lượng video ở đây trong QoE (**chất lượng video vẫn không thay đổi trong quá trình phát trực tuyến video vì Douyin đã áp dụng chiến lược tải xuống liên tục**)**

**Text

Description automatically generated**

* **Cost: wt** là nội dung được lưu vào bộ đệm nhưng chưa bao giờ được xem do các sự kiện cuộn sớm

**A picture containing text, clock, gauge

Description automatically generated**

* **wt > 0:** khi Edt khác rỗng
* **Loss function: kết hợp QoE và Cost**
* A picture containing clock, watch

  Description automatically generated
* n: weight kiểm soát sự cân bằng giữa thời gian dành cho rebuffering và thời gian dành cho việc down video không sử dụng

Mục đích: tìm π (a | s), nó cho agent biết xác suất mà một hành động sẽ được chọn với trạng thái hiện tại

Optimize:

A picture containing text, clock

Description automatically generated