

Вступ

Однією з найбільш широкоживаних технологій сучасності є відеострімінг. Це саме та технологія, завдяки якій ми маємо численні додатки для відеодзвінків, відеоконференцій, вебінарів, цифрове телебачення, віддалене керування дронами, розумні камери відеоспостереження і розумні, відеодзвінки тощо.

В основі технології лежить кодування відеопотоку і його передача мережею зі сторони передавача, та приймання і розкодування на стороні приймача. На практиці, реалізація даної технології стикається з проблемами непостійності доступної пропускної здатності мережі та втратами фрагментів відеопотоку. Своєю чергою, відеопотік також має змінну кількість бітів даних за одиницю часу залежно від характеру інформації. Реалізація даної технології повинна ефективно реагувати на зміни в мережі та в самому відеопотоці, маючи при цьому обмежені можливості в оцінці пропускної здатності мережі та розміру відеопотоку в наступний момент часу. Таким чином, динамічна оцінка пропускної здатності мережі, реакція на зміни, її швидкість і коректність адаптації до нових умов і наявних ресурсних обмежень є питаннями, які стали предметом наукових і практичних досліджень.

Принципова схема технології

Передавальна частина більшості рішень в галузі відеострімінгу можна представити наступною композицією функцій:

Передати(Закодувати((ЗгенеруватиКадри()))

І зі сторони отримувача:

Відтворити(Розкодувати((ОтриматиКадри()))

Де,

ЗгенеруватиКадри – функція генерації відеокадрів. Це може бути або захоплення відеокадрів напряду з пристрою камери або шляхом читання даних з носія або БД.

Закодувати – з метою зменшення розміру потоку або зменшення ресурсів потрібних для обробки, відеокадри кодуються спеціалізованими кодеками які часто продукують відеопотік змінного об'єму за одиницю часу. Огляд принципів домінуючого способу кодування H.264 можна знайти в [3].

Розкодувати – зворотня функція до кодування.

Передати – передача закодованої інформації до споживача або проміжного агента по мережі. В якості протоколу транспортного рівня можуть бути різні варіанти в залежно від застосування.

ОтриматиКадри – зворотня функція до передачі. Отримання з метою реконструкції у відеопотік належної якості.

Оптимізація кожної з цих функцій окремо як і композиція яка реалізує систему в цілому є предметом окремих досліджень серед науковців і практиків. Додатковий вимір в дослідженнях створює різномайття застосувань технології та обмеження в доступних обчислювальних ресурсах. В загальному випадку сформулювати набір критеріїв і метрик не є можливим, а все визначається специфікою застосування. Наприклад, стрімінг відео великої роздільної здатності між сучасними персональними комп'ютерами в мережі Інтернет може потребувати інших підходів ніж передача відеопотоку з дистанційно керованого дрона в локальній безпроводній мережі. Наявні обчислювальні ресурси теж можуть накладати свої обмеження. Деякі фактори які визначають відмінності:

1. Бажані затримки на початок трансляції і інші аспекти ініціації сесії. Може варіюватися від сотень мілісекунд до десятків секунд залежно від області застосування. Існує чимало протоколів ініціації і керуванням сесії, проте з новими застосуваннями може виникати необхідність в проєктуванні нових, здатних задовольняти специфічним функціональним або нефункціональним потребам.
2. Актуальність відеопотоку. В таких застосуваннях як відеодзвінок чи керування дроном отримувати відео з затримкою в більше ніж, скажімо, 2 секунди може бути неприпустимим, а від так, система має мати принципово інший підхід до адаптації ніж, скажімо, стрімінг відеофільму в моделі *на-вимогу*.
3. Пріоритет якості зображення. Залежно від типу застосування й обмежень в пропускній здатності каналу передачі, пріоритет може надаватися тим або іншим параметрам. Для одних застосунків важлива максимальна роздільна здатність і відсутність небажаних артефактів, в інших це може бути максимальна частота кадрів. Бувають також і гібридні схеми, коли окремі зони в кадрі кодуються з меншими втратами якості ніж інші. Приклад керування потоком з врахуванням регіонів можна знайти в [8]. Основними, найбільш

визначними параметрами виступають кількість кадрів в секунду та роздільна здатність.

4. Швидкість адаптації до мережевих обмежень. В переважній більшості застосувань, включаючи потокове відео на вимогу, при зміні можливості пропускної здатності мережі система повинна підлаштуватися до нових умов якнайшвидше і в задовільний для даного застосування спосіб. Якщо пропускна здатність каналу не є постійною, то модуль передачі може стикнутися з неможливістю передавати відеопотік у всьому обсязі за одиницю часу. Втім, принципово, система може підлаштуватися і зменшити або роздільну здатність, або частоту кадрів, або ж змінити параметри кодування на такі, які зменшать розмір ціною зменшення якості зображення. Для цього системі потрібен ефективний алгоритм здатний адаптивно змінювати параметри зображення і необхідні ретрансляції якщо мова йде про канали передачі без гарантій гарантованого доставлення даних. Окремою задачею є постійна оцінка пропускної здатності мережі в реальному часі, що для складних багаторангових мереж типу Internet є самостійною гілкою досліджень.

Для побудови ефективних застосувань технології відеострімінгу, концептуально обчислення мають бути дещо складніші. А саме, функція передавання в якості аргументів приймає свій стан і наряду з параметрами енкодера продукує свій наступний стан на основі статистичної інформації про процес самої передачі:

$$\text{Передати}(V, S, N, F) \rightarrow \{S', K\}$$

Де, V – відеопотік, S – стан обчислень, S' – наступний стан, N – метадані про процес передачі, F – дані отримані від приймача, K – параметри кодування для енкодера.

Для застосувань які вимагають високої актуальності відеопотоку, функція передавання може базуватися на протоколі передачі даних без гарантованого доставлення даних і повинна сама займатися пересиланнями, що створює додатковий простір для алгоритмічних оптимізацій, звідси і параметр F , який представляє метадані про процес отримання на стороні приймача. В свою чергу, функція кодування має додатковий аргумент який прямо або непрямо керує динамічним розміром відеопотоку:

$$\text{Закодувати}(V, K) \rightarrow V'$$

Де, V – вхідний відеопотік, K – параметри кодування, V' – вихідний відеопотік готовий до передачі.

В такому виді це створює систему зі зворотним зв'язком, де в процесі передавання даних ми отримуємо інформацію про стан процесу передавання і приймання і використовуємо цю інформацію для зміни параметрів кодування з метою забезпечення оптимальної для даного застосування продуктивності. Погляд на дану проблематику як на систему керування з пропорційно-інтегрально-диференціальним законом керування можна знайти в [4].

Критерії оцінки продуктивності

Говорячи про оптимальну продуктивність для того чи іншого застосування, слід сформулювати чіткі критерії оцінки. Існує чимало загальних методів для оцінки якості відеострімінгу. Оскільки суть даної технології це передача відео для споживання людиною, то й відправною точкою в оцінці якості теж є суб'єктивна оцінка сприйняття людиною. На практиці ж, для оцінки продуктивності рішень використовуються синтетичні метрики. Вимірювання якості зображення є старою проблемою до якої було запропоновано чимало простих і практичних підходів. Mean-square-error (MSE), Peak-Signal-to-noise-ratio (PSNR) і Structural Similarity Index (SSIM) є прикладами таких метрик які були придумані для статичних зображень, а пізніше адаптовані до відео. Існують також дослідження і комерційні моделі на базі таких досліджень, які базуються на машинному навчанні та здатні надавати оцінки які дуже добре корелюють з людським сприйняттям.

З розвитком машинного навчання з'явилася можливість оцінювати якість відео в межах роботи самого протоколу, в реальному часі, та використовувати нейронні мережі. Короткий огляд методів оцінки адаптивного стрімінгу можна знайти в [7]. Слід зауважити, що такі методи не можуть бути доступними для всіх застосувань в силу підвищених обчислювальних ресурсів притаманних для систем на базі нейронних мереж.

Крім метрик метою яких є висока кореляція з суб'єктивним сприйняттям людиною, можуть існувати більш технічні метрики які важливі для специфічних застосувань. Наприклад, для такого застосування як дистанційне керування роботизованим пристроєм, за умов зменшення пропускну здатності каналу, продовжити показувати відеопотік з мінімальними затримками на передавання та отримання, навіть з істотними втратами якості може бути прийнятним. Всі ці критерії є специфічними для застосування і визначають цільову функцію яку алгоритм передавання намагається максимізувати.

Огляд публікацій

Даний розділ містить далеко не виключний перелік публікацій які можна знайти з допомогою пошукових систем за ключовими словами тих чи інших відомих методів які мають місце в технології. Подальші дослідження в області вимагатимуть значно глибшого ознайомлення найбільш цитованих публікацій релевантних до вибраної предметної області.

В [1] дослідники пропонують новітню систему оцінки пропускну здатності мережі на базі методів глибинного навчання з підкріпленням. Хоча цей спосіб може бути незастосовним для застосувань для вбудованих пристроїв, все ж це може мати місце в якості допоміжного інструменту або як еталон для порівняння з простішими методами.

В [2] можна знайти огляд ключових ідей і класифікацію методів систем з адаптивним розміром відеопотоку (*ABR — adaptive bitrate*), що є ключовим елементом систем для передачі відео в реальному часі в мережах загального призначення.

В [3] авторами пропонується сучасний, просунутий протокол з підтримкою *ABR*, на базі базового транспортного протоколу без гарантованої доставки. Дослідники не лише зосереджуються на блоці керування передаванням, але й експлуатують особливості кодеків нового покоління.

В [4] можна знайти погляд на проблему, як на задачу з області систем керування, яка відома наявністю теоретичної бази і засобами моделювання.

В [5] автори проводять короткий огляд актуальних технологій відеострімінгу, їх класифікацію і метрики для оцінювання.

В [6] можна знайти огляд практичної оцінки якості відеострімінгу від лідера в області розповсюдження відео контенту Netflix.

В [7] визначається поняття якості досвіду (*QoE – quality of experience*) і проводить короткий огляд методів оцінки якості адаптивного відеострімінгу.

В [8] пропонується метод використання так званих регіонів інтересу (*ROI – region of interest*) в застосуванні летючих дронів, що може бути релевантним до класу відеострімінгу реального часу для вбудованих пристроїв.

План діяльності

В своїй аспірантській діяльності я б хотів займатися пошуком нових, або вдосконаленням існуючих керуючих алгоритмів адаптивного відеострімінгу реального часу. Це може бути новий керуючий алгоритм для керування передаванням, або вдосконалення існуючого, який би відкривав нові можливості для застосування в якійсь специфічній області застосування. Можливо, покращенням існуючих допоміжних алгоритмів, або ж пошуком комбінації існуючих алгоритмів і методів для досягнення нових, раніше недоступних можливостей.

Перший рік я б хотів присвятити огляду найбільш цитованих релевантних публікацій в даній області. А також, обов'язковим кроком є розробка базового рішення яке б реалізувало основні блоки тих аспектів технології, які будуть предметом подальших досліджень. Це дозволить глибоко зануритися в предметну область, зрозуміти проблематику, і що не менш важливо, це зможе стати основою для тестування майбутніх ідей і оцінки їх продуктивності. Таке рішення повинно бути структуровано таким чином, щоб воно дозволило розвивати окремі блоки залежно від пріоритетів в дослідженнях. Також, така тестова система має дозволяти отримувати кількісні і якісні характеристики продуктивності.

Проаналізувавши релевантні публікації та маючи базову реалізацію на якій можна проводити експерименти, далі можна буде зосередитися на тому чи іншому напрямку досліджень. Актуальним і в меншій мірі вивченим може бути така область як відеострімінг реального часу з можливістю для застосування у вбудованих пристроях. Притаманними характеристиками таких систем є потреба у швидкій ініціації, високій актуальності відеопотоку що визначається низькими порогами затримок між передаванням і отриманням кадрів, специфічними вимогами до якості зображення на стороні приймача та обмежені обчислювальні ресурси.

Посилання

1. ACM MMSys 2024 Bandwidth Estimation in Real Time Communications Challenge: <https://arxiv.org/html/2403.06324v2>
2. Adaptive Bitrate Video Streaming for Wireless nodes: A Survey: https://www.researchgate.net/publication/343415080_Adaptive_Bitrate_Video_Streaming_for_Wireless_nodes_A_Survey
3. Adaptive Scalable Video Streaming (ASViS): An Advanced ABR Transmission Protocol for Optimal Video Quality: <https://www.mdpi.com/2079-9292/12/21/4542>
4. A control theoretic approach to ABR video streaming: A fresh look at PID-based rate adaptation: <https://ieeexplore.ieee.org/document/8057056>
5. Video Quality Metrics: https://www.zte.com.cn/content/dam/zte-site/res-www-zte-com-cn/mediarees/magazine/publication/com_en/article/201901/en201901003.pdf
6. Toward A Practical Perceptual Video Quality Metric: <https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652>
7. A Brief Survey on Adaptive Video Streaming Quality Assessment: <https://arxiv.org/abs/2202.12987>
8. Region of Interest (ROI) Coding for Aerial Surveillance Video using AVC & HEVC: <https://arxiv.org/abs/1801.06442>