Distribution of News Videos to Mobile Devices over Challenged Networks

Shu-Ting Wang
Outline

- Introduction
- System Overview
- Distribution Planning Problem
- Trace-driven Simulation
- System Prototype
- Conclusion
Background

- Mobile devices are getting increasingly popular in developing countries, where mobile users rarely have the Internet access.
Motivations

- Conquer digital divide in rural area with weak network infrastructures for online services
- Online banking, learning, health service
- Online news reports from CNN, BBC, Al Jazeera, and etc
What’s Challenged Networks

- Intermittent connectivity
- Scarce resources
- Widely varying network conditions
A User Story

Watch at home

Challenged CDN

Get Video 3

Ad-hoc networks

Share Video 2

Local Proxy: caches multimedia content

School

City hall

Distribution Server

Multimedia Content Network

Receive multimedia content from proxies

Video 1

Video 2
What’s Challenged CDN?

- Challenged Content Delivery Networks (CCDNs) help news providers reach out to citizens living in rural areas without the Internet access.
Outline

- Introduction
- **System Overview**
- Distribution Planning Problem
- Trace-driven Simulation
- System Prototype
- Conclusion
Objective

- The CCDN is aim to create a distribution plan, in order to intelligently distribute news report
  - The best time
  - The right mobile users
  - The highest possible quality
System Architecture (1/2)

New Reports

Preferred Keywords

User Profiles

Contact History

Keyword Extractor

Keywords

Israel continues its ...... on Gaza

Viewing Prob. Ranker

News Ranking

User Contact Predictor

Future Contacts

Distribution Planning Algorithm

Distribution Plan
System Architecture (2/2)

Distribution Server

- New Reports
- Preferred Keywords
- User Profiles
- Contact History

Keywords: Israel continues its ...... on Gaza

- Viewing Prob. Ranker
- News Ranking
- User Contact Predictor
- Future Contacts

Distribution Planning Algorithm

Local Proxy

- User Matcher
- Distribution Plan

Phones with location and user profile icons.
System Models

● Network Model
  ○ A content network that delivers news video
  ○ Challenged networks where mobile users rarely have Internet access

● News Model
  ○ Multi-layer representation of news

● Mobile Users Model
  ○ Users’ interest and mobility are predictable
  ○ 85% of time a mobile users stays at his/her top 5 favorite locations
News Model

- A news report contains multiple layers: texts, audio, and layered video
- Its layers have different improvement of user experience
- To match against mobile user interests, each news reports has one or multiple topics

Israel continues its …… on Gaza
Mobile Users Model

- A mobile user has
  - An interested keywords set
  - A trajectories set recorded series of timestamped locations of the user
Distribution of News Reports

- Our distribution planning algorithm computes the distribution plans for all users.
  - The distribution server pushes the distribution plan and user profile to the local proxies
  - The mobile user fetches the distribution plan when he’s close to particular local proxies
  - Mobile users exchange their distribution plans based on their interests
Outline

- Introduction
- System Overview
- Distribution Planning Problem
- Trace-driven Simulation
- System Prototype
- Conclusion
Problem Formulation: Objective

Maximizes the expected overall user experience across all mobile users.

Decision variable for:
1. which contact with which user to communicate
2. which news and layer of news to fetch

User experience improvement on the accumulated layers of news

User’s viewing probability of a news report
Problem Formulation: Constraints (1/4)

\[ \text{st : } \psi_{p_{u',c'},n'} \geq \psi x_{u',n',v',c'} \quad \forall 1 \leq u' \leq U, 1 \leq n' \leq N, 1 \leq v' \leq L, 1 \leq c' \leq C \]

Mobile users never request a news report from someone who is unlikely to watch it

\[ \sum_{n=1}^{N} \sum_{l=1}^{L} b_{nL+l} x_{u',n,l,c'} \leq r_{u',c'} K_{u',c'} \quad \forall 1 \leq c' \leq C, 1 \leq u' \leq U \]

Each contact duration is long enough for the planned unit transfer under the given transmission throughput
Problem Formulation: Constraints (2/4)

\[
\sum_{c=1}^{C} x_{u',n',l',c} \geq \sum_{c=1}^{C} x_{u',n',l'',c} \quad \forall 1 \leq u' \leq U, \ 1 \leq n' \leq N, \ 1 \leq l' < l'' \leq L
\]

Layer dependency, i.e., higher layer is only decodable when all lower layers are received

\[
\sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{c=1}^{C} b_{nL+l} x_{u',n,l,c} \leq d_{u'} \quad \forall 1 \leq u' \leq U
\]

The disk budget for user u’

Total size of planned new videos does not exceed the disk budget
Problem Formulation: Constraints (3/4)

\[
\sum_{u=1}^{U} \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{c=1}^{C} \left\{ 1 - \min \left[ 1, \max (p_{u,c}, u') - \min (p_{u,c}, u') \right] \right\} b_{nL+l} x_{u,n,l,c} \hat{e}_{u',c} 
\]

This is 1 iff \( p_{u,c} = u' \), then the mobile user \( u \) plans to receive news videos from mobile user \( u' \)

\[
+ \sum_{n=1}^{N} \sum_{l=1}^{L} \sum_{c=1}^{C} b_{nL+l} x_{u',n,l,c} \hat{e}_{u',c} \leq q_{u'} \quad \forall 1 \leq u' \leq U
\]

transmission per-unit energy consumption

The energy budget for user \( u' \)

The total transmitting and receiving energy do not exceed the energy budget
Problem Formulation: Constraints (4/4)

\[ \sum_{c=1}^{C} x_{u',n',l',c} \leq 1 \quad \forall 1 \leq u' \leq U, 1 \leq n' \leq N, l \leq l' \leq L \]

Each user does not receive the same unit multiple times

\[ x_{u,n,l,c} \in \{0, 1\} \quad \forall u, n, l, c \]

\[ X_{u,n,l,c} \] is a 0-1 decision variable
Multidimensional Knapsack Problem

- Our problem is a 0-1 Integer Linear Programming problem.
- Our formulation can be transformed into a 0-1 Multidimensional Knapsack Problem (MKP).

\[
\begin{align*}
\text{max} & \quad \sum_{j=1}^{J} r_j x_j \\
\text{s.t.} & \quad \sum_{j=1}^{J} w_{k,j} x_j \leq y_k \quad \forall k = 1, 2, \ldots, K
\end{align*}
\]

- \( x_j \in \{0, 1\} \quad \forall j = 1, 2, \ldots, J. \)
- \( r_j \) stands for the profit of having object \( j \)
- \( w_{k,j} \) is the resource consumption for \( k \)th constraint.
Problem Transformation

• Set $J = UNLC, \quad r_j = \rho_{nL+l\psi u}, n$

\[ \max \sum_{j=1}^{UNLC} (\rho_{nL+l\psi u}, n)_j x_j \]

• Then, where $j = uNLC + nLC + lC + c \ \forall u, n, l, c$

• Next, move the decision variable to the left hand side, and leave others to the right hand side
MKP can be solved by several exact and heuristic algorithms proposed in the literature.

We leverage CPLEX, a commercial solver.

The workflow of our proposed Distribution Planning Algorithm:

- Transformation the news distribution problem into an MKP
- Solving MKP using CPLEX with heuristics
- Converting the MKP solution into a distribution plan
Adaptive Communication Strategies (1/2)

- A mobile device to the local proxy
  - If we do not have a plan, get one
  - Follow the plan to retrieve planned units for this contact on the local proxy
  - If finished, retrieve planned units for other contacts
  - Once the plan is completed, the local proxy runs recommendations for more news based on the user’s interest
Adaptive Communication Strategies (2/2)

- A mobile device to a mobile device
  - Exchange the units that are planned to request from each other in this contact
  - If finished, exchange planned units for other contacts
  - Once the plan is completed, make the other mobile device to select particular news

- The system prefer units that have higher ratio of user experience improvement normalized to unit size
Outline

- Introduction
- System Overview
- Distribution Planning Problem
- Trace-driven Simulation
- System Prototype
- Conclusion
Datasets

● User contacts:
  ○ Microsoft GeoLife dataset: GPS trajectories of about 180 mobile users over 4 years
  ○ SIGCOMM dataset: 76 users of bluetooth contact over 3 days

● Video reports: 27 news reports from Al Jazeera in mid-July 2014, and divide each news report into five layers

● User interests: leveraging the user queries in the Microsoft LETOR dataset
Performance Metrics - User

● User experience
  ○ the summation of the user experience of all the watched news reports

● Viewing ratio
  ○ the fraction of viewed news reports among all the downloaded ones

● Missed ratio
  ○ the fraction of unavailable news reports among all the user demanded
Performance Metrics - System

● Energy consumption
  ○ the energy consumption of mobile devices

● System efficiency
  ○ the ratio of user experience and energy consumption

● Used disk space
  ○ the amount of used disk space.
Simulation Setup

- Disk Budget: 60, 125, 250, 500 MB
- Energy Budget: 500, 1000, 2000, 4000 J
- WiFi power consumption: 600 mW in the range of 25 m
- The maximal segment size is 25 MB
- We set the user experience improvements of layers 1–3 to be 0.26, and the medium- and high-resolution videos to be 0.12 and 0.10
Baselines of CCDN Simulator

- Experience driven baseline (Base$_v$)
  - Greedily requests the unit with the \textit{highest user experience improvement}

- User centrality driven baseline (Base$_u$)
  - Greedily sends the units to the mobile device with the \textit{highest number of contacts}

- Throughput driven baseline (Base$_n$)
  - Greedily sends the units to the mobile device with the \textit{best channel condition}
Service Quality: User Experience

- CCDN outperforms 30% of Base_v, and overwhelms the other baselines
- CCDN considers more improvement of user experience compared to Base_v
Service Quality: Missed Ratio

- The mobile users miss about 80% of new if $\text{Base}_u$ or $\text{Base}_n$ are used.

- CCDN has only 50% missed ratio and $\text{Base}_v$ is at most 10% worse than CCDN.
Service Quality: SIGCOMM dataset

- The SIGCOMM dataset is so well-connected, user receive at least mid-resolution
- It leads to no optimization room for CCDN
System Efficiency

- More than 10 times improvement compared to \( \text{Base}_u \) and \( \text{Base}_n \)
- Up to 25% of improvement compared to \( \text{Base}_v \)
Resource Efficiency: Disk Space

- Our algorithm uses at most 1/3 of disk space, compared to all three baseline algorithms.
Resource Efficiency: Viewing Ratio

- Our CCDN wastes the least downloaded units
- Base_u and Base_n achieves less than 50% of viewing ratio of CCDN
- Base_v is 5% lower than CCDN
Different Energy Budgets

- Our CCDN has the 5 times user experience gain and 50% miss ratio reduction.
Different Disk Budgets

- Our CCDN achieves roughly the same service quality with growth of disk budget.
- CCDN uses disk budget efficiently.
Running Time of CCDN

- The daily running times are 12 mins and 100 mins
- ILP solver runs fast, while ranker takes time
Machine Learning Algorithms

- Adaptive communication strategies work with user experience improvement.
- The ranker requires fine tuning on parameters.
Outline

● Introduction
● System Overview
● Distribution Planning Problem
● Trace-driven Simulation
● System Prototype
● Conclusion
System Prototype Architecture

Local Proxy

- dhcpd
- apache
- hostapd

Download Plans and News

Upload User Profiles

Distribution Server

- Distribution Algorithm
- User Profiles
- News
- NBC News Downloader

Plan

User Profiles

News
Experiment Setup

- **Resource budget:**
  - Disk: 2GB, 25% of the disk space
  - Energy budget: 6000 J 20% of the battery

- A distribution server and two local proxies are all on campus

- The distribution server sends the latest 300 news reports to local proxies everyday. Each news report has 5 layers, so we have 1500 units in total

- Mobile users refresh their news storage 5 a. m. everyday
Resource Consumption

- Diverse network condition leads to different resource consumption
Contact Duration and Remaining Units

- Shorter contact duration leads to more remaining units
Prototype Limitations

- No units come from smartphone peer
- Large videos require segmentation for efficient download
- A machine learning algorithm with appropriate parameters to support better adaptation among diverse users
- Make distribution server and local proxies may be in different geographic locations with diverse network conditions
Outline

● Introduction
● System Overview
● Distribution Planning Problem
● Trace-driven Simulation
● System Prototype
● Conclusion
Conclusion

- We proposed a Challenged Content Delivery Network (CCDN), which carefully plans the distribution of news reports to mobile users.
- CCDN outperforms the baseline algorithms by 55% to 10 times of user experience.
- Achieves higher system efficiency than the baseline algorithms by 37% to 20 times.
- Fast running time in 12 mins for 150 users.