An Efficient Content Delivery Infrastructure Leveraging the Public Transportation Network

Qiankun Su, Katia Jaffrès-Runser, Gentian Jakllari and Charly Poulliat

IRIT, INP-ENSEEIHT, University of Toulouse

Nov. 16th, 2016

CNRS - INPT - UPS - UT1 - UTM

Problem statement and motivation

Growing urbanization

- 70% of the world's population lives in urban areas by 2050[1].
- Growing urbanization leads to an increasing demand for public transport.
- A significant part of mobile contents is consumed while people use public transportation.

Growing mobile data traffic

- The mobile data traffic will be increased nearly 8-fold between 2015 and 2020[2].
- 4G can not satisfy the fast-growing demands.
- 5G deployments are not expected until at least 2020[2].

[1] *United Nations, World Urbanization Prospects: The 2014 Revision*. https://esa.un.org/unpd/wup. URL: https://esa.un.org/unpd/wup/.

[2] Visual Networking Index Cisco. "Global mobile data traffic forecast update, 2015–2020". In: *white paper* (2016).

Contributions

A novel content delivery infrastructure

- relieve the wireless bandwidth crunch in urban centers
- offload up to 1TByte within 12 hours[3]

[3] Transmission rate: 100Mb/s

Outline

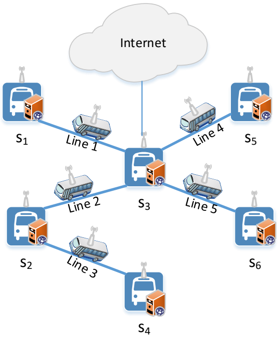
Outline

1. A content delivery infrastructure using PTNs
   - Scenario descriptions
   - Our routing policy
2. XOR network coding for PTNs
   - Problem statement and motivation
   - XOR network coding implementation
   - Performance evaluation
3. A cost-effective and secure design
   - Problem statement and motivation
   - 3-Tier architecture
     - 1-Tier
     - 2-Tier
     - 3-Tier
   - Cost-effectiveness analysis

Scenario descriptions

Scenario descriptions

- Public buses act as data mules, creating a delay tolerant network (DTN)
- Contents are obtained from nodes connected to the Internet
- Passengers download/upload contents from/to buses or bus stations



Where to install WiFi AP?

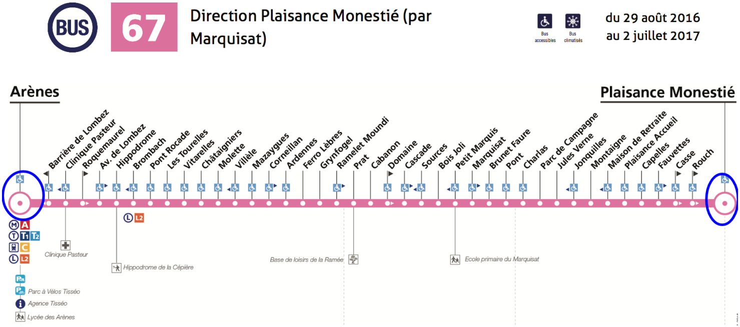


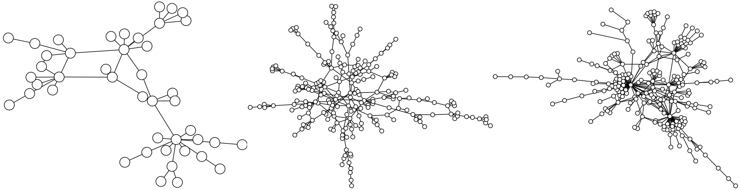
Fig. 2: The bus line 67 in Toulouse, France

- The waiting time of buses at intermediate stops is very short.
- Wireless access points (APs) are deployed at end stations, but not intermediate stops.

Modeling

Our infrastructure can be modeled as an undirected graph where

- nodes represent end stations
- edges represent bus lines



(a) Toulouse (b) Paris (c) Helsinki

Fig. 3: The biggest connected component of public transportation networks.

Our routing policy (1/2)

Main routing protocols in DTN are designed for

- non-predictable mobility patterns

The features of DTNs created by PTNs,

- The network topology is stable.
- The behavior of buses is predictable.

Our routing policy,

- messages are delivered following the shortest path
- pre-calculate routing tables for each end station

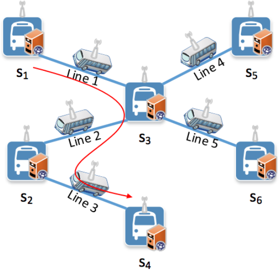


Fig. 4: Content delivery using PTNs.

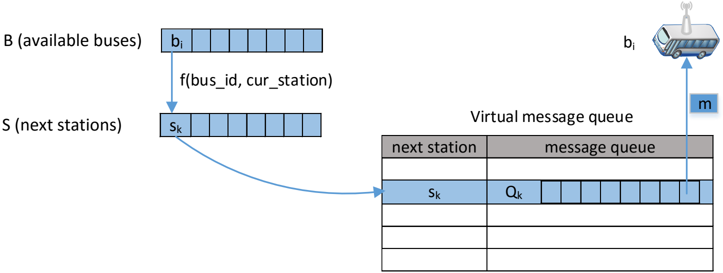
Our routing policy (2/2)

Receive a message m at an end station

- extract m's destination, look up its next-hop stations s_k
- m is placed into Q_k that stores messages going to s_k

Send a message m at an end station

- B: a list of buses currently waiting at the station
- S: a corresponding list of next-hop stations



Outline

Problem statement and motivation

- PTNs are built around the concept of hubs with many bus lines.
- The fair medium access control.

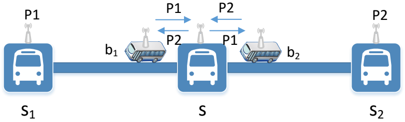


Fig. 5: Light network load

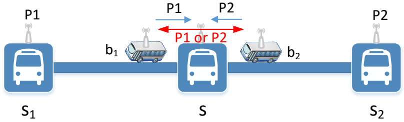


Fig. 6: Heavy network load

$\Rightarrow$ Such an imbalance results in a significant drop in throughput under heavy traffic conditions.

XOR network coding

XOR network coding implementation

- pairwise inter-session flows[4]
- hop-by-hop

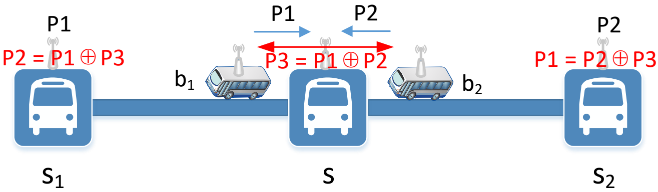


Fig. 7: The benefits of XOR network coding.

[4] Q. Su et al. "XOR network coding for data mule delay tolerant networks". In: *2015 IEEE/CIC International Conference on Communications in China (ICCC)*. 2015, pp. 1–6. DOI: 10.1109/ICCChina.2015.7448634.

XOR network coding implementation

Encoding procedures:

- Message queues Q_{ij} are indexed by the previous station s_i and the next station s_j of messages.

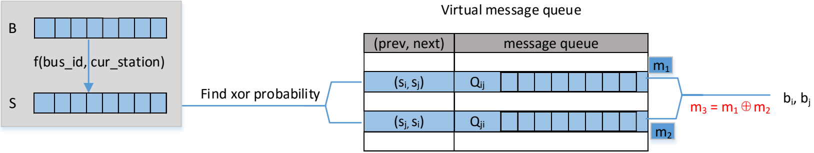


Fig. 8: Encoding procedures

Decoding procedures:

- receive a xor-ed message $m = m_i \oplus m_j$
- xor-ing again with the message previously sent, $m_j = m \oplus m_i$

Simulation setup

This paper leaves for further investigation how content is requested, updated and fetched. The goal is to show the pure network coding benefits in our infrastructure.

Simulator

- the ONE (Opportunistic Network Environment simulator)

Mobility model

- Real traces in GTFS [5] from Toulouse, Paris and Helsinki, $7:00$ to $19:00$

Data flows (multiple unicast flows)

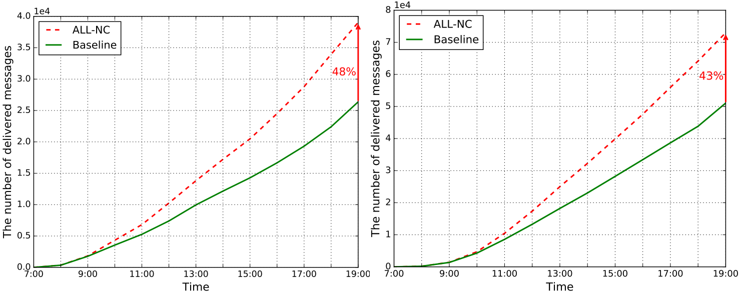
- A message is created at every end station with a given creation period ($\Delta = 20$ *seconds*).
- The message destination is selected uniformly at random among all the stations.

[5] GTFS (General Transit Feed Specification), developed by Google, is a common format for public transportation schedules and associated geographic information.

Performance evaluation

Benefits of network coding,

- ALL-NC: with network coding
- Baseline: without network coding
- X axis: the period from 7:00 to 19:00
- Y axis: the number of delivered messages



(a) Toulouse (b) Paris

Fig. 9: Network coding benefit : number of delivered messages.

Outline

Problem statement

Before adapting network coding as part of our infrastructure, two main challenges have to be addressed.

- Network coding is threatened by pollution attack and a *secure* network coding AP is more expensive than a network coding AP.
- Installing network coding enabled APs at all stations is expensive, not to mention *secure* network coding APs.

3-Tier architecture

Divide stations into 3 tiers,

- No AP is deployed (1-Tier)
- Regular AP is deployed (2-Tier)
- Secure network coding enabled AP is deployed (3-Tier)

The goal is to,

- guarantee end-to-end delivery
- minimize the cost of deployment

1-Tier

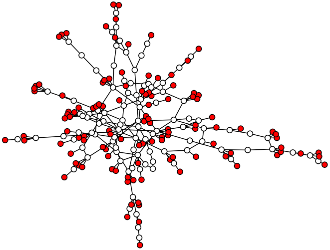


Fig. 10: The leaf nodes in Paris topology (in red)

Leaf station removal saves 63.5%, 53% and 52% of wireless APs.

City	Baseline	ALL-NC	Leaf stations removal
Toulouse	44	44	16
Paris	213	213	99
Helsinki	217	217	90

Table 1: Number of wireless access points required to cover 3 different cities.

2-Tier

A minimum connected dominating sets,

- minimize the number of wireless AP
- guarantee the end to end connectivity

A CDS is formed by *M Rai et al.*[6].

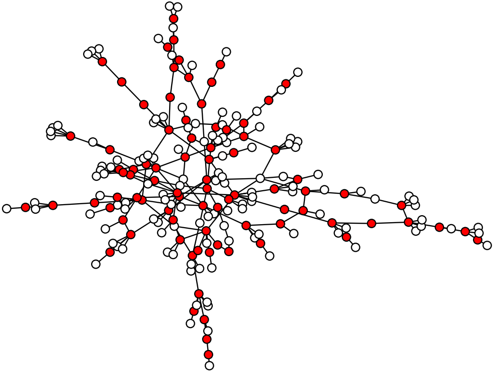


Fig. 11: A CDS in Paris topology (in red).

2-Tier, the decrease of APs

Save around three times of wireless APs.

City	Baseline	ALL-NC	2-Tier
Toulouse	44	44	13
Paris	213	213	85
Helsinki	217	217	60

Table 2: The 2-Tier architecture reduces the required number of interfaces by approximately a factor of 3.

2-Tier, performance evaluation

- 2-Tier: stations belong to CDS equipped with network coding enabled APs

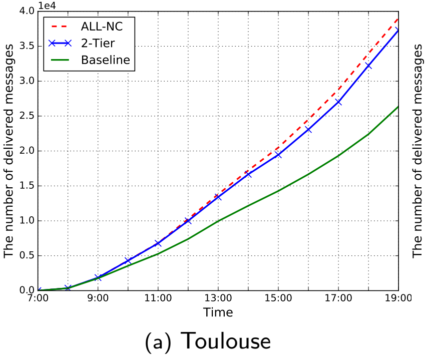


Fig. 12: Number of messages delivered for Baseline, ALL-NC and 2-Tier.

3-Tier

Select the top *n* nodes from CDS to install network coding AP

- Large benefits of network coding if existing a lot of cross flows
- Identify nodes with high degree, betweenness, PageRank

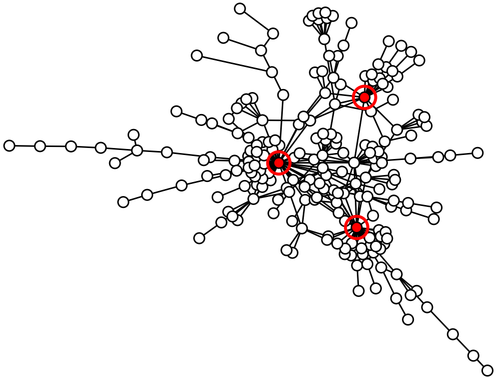
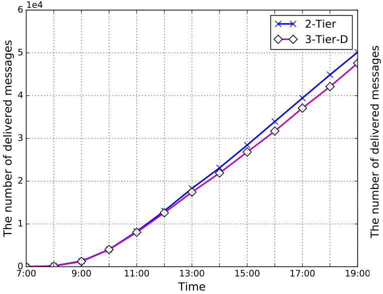


Fig. 13: The top 3 highest PageRank in Helsinki topology (in red)

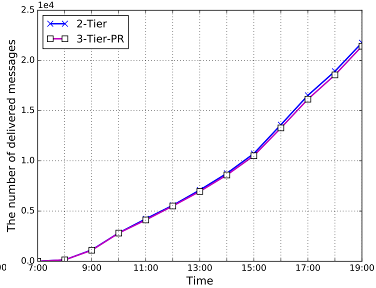
3-Tier, performance evaluation

City	2-Tier	3-Tier	Metric
Toulouse	13	2	Degree
Paris	85	10	Degree
Helsinki	60	3	PageRank

Table 3: 3-Tier reduces the number of such interfaces by over an order of magnitude.



(a) Paris



(b) Helsinki

Fig. 14: Packets delivered for 2-Tier and 3-Tier.

Cost-effectiveness analysis

The cost of

- a regular wireless AP: 1
- a secure network coding enabled AP: C $(C > 1)$
- Y axis: the cost effectiveness $= \frac{The\ number\ of\ delivered\ messages}{The\ deployment\ cost}$

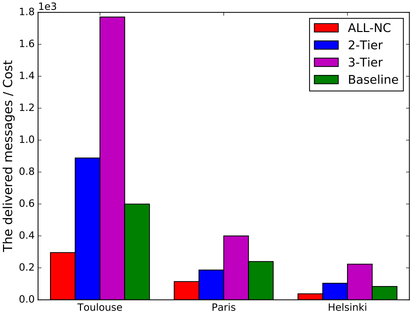


Fig. 15: The cost effectiveness for all architectures ($C = 3$).

Conclusion

A cost effective content delivery infrastructure (a 3-Tier architecture)

- guarantees end-to-end connectivity
- provides high packet delivery
- minimizes deployment cost

Real trace-based simulation

- reduce the number of wireless APs by a factor of 3
- deliver more messages than a baseline architecture
- offload a large amount of data, e.g., 1TByte within 12 hours in Paris topology

Thank you for your attention.

Q & A

qiankun.su@enseeiht.fr