Not All Content is Created Equal: Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination

Pavlos Sermpezis   Thrasyvoulos Spyropoulos

Mobile Communications Dept.
EURECOM
Sophia-Antipolis, France
Technologies and Motivation

• Opportunistic / D2D Network
  - Mobile Devices: smartphones, laptops, etc.
  - Direct communications: Bluetooth, WiFi direct, etc.

• Mobile Data Offloading
  - Fixed network, femtocells, etc.
  - Opportunistic Networking
  - D2D communications
Cellular and D2D Network

- Holders
- Requesters

Not All Content is Created Equal: Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination.
Availability and Popularity

- **Availability**: #Holders
- **Popularity**: #Requesters
**GOAL 1**
Given:
- content availability (storage) & popularity
- nodes cooperation (no network control) & mobility
  → *analyze performance*

**GOAL 2**
Given:
- a performance scenario (mobility, popularity, etc.)
  → *optimize storage* (i.e. availability / holders allocation)
Mobility

- $t_x$: random process with rate $\lambda_{ij}$ (for a pair \{i,j\})
- $\Delta t$: exponential, Pareto, etc.

Content Traffic

- Popularity:
  \[ P\{ \text{#requesters} = n \} = P_p(n) \]
- Availability:
  \[ P\{ \text{#holders} = m \mid \text{#requesters} = n \} = g(m \mid n) \]
Goal 1: Performance Prediction

Content Access Time $T_M$:
- depends on mobility ($\lambda_{ij}$), #holders ($m$), #requesters ($n$)

Generic results:

$$E[T_M] = \frac{1}{E_p[n]} \cdot E_p \left[ n \cdot \sum_{m} E_{m\lambda} \left[ \frac{1}{\lambda_{ij}} \cdot \alpha(m|n) \right] \right]$$

$$P\{T_M \leq TTL\} = 1 - \frac{E_p \left[ n \cdot \sum_{m} E_{m\lambda} \left[ e^{-x \cdot TTL} \cdot g(m|n) \right] \right]}{E_p[n]}$$

Extension: Availability - popularity correlation
Goal 1: Performance Prediction

Content Access Time $T_M$:
- depends on mobility ($\lambda_{ij}$), #holders ($m$), #requesters ($n$)

Example
Popularity: Pareto ($a=2, n_0$)
Mobility: $\lambda_{ij}$ gamma distributed ($\mu_\lambda, CV_\lambda$)
Availability: $g(m,n) \equiv g(n) = c \cdot n$

$$E[T_M] = \frac{1}{\mu_\lambda \cdot CV_\lambda^2} \left[ \frac{c \cdot n_0}{CV_\lambda} \cdot \ln \left( \frac{1}{CV_\lambda^2 \cdot c \cdot n_0} \right) - 1 \right]$$
Goal 1: Performance Prediction

Content Access Time $T_M$:
- depends on mobility ($\lambda_{ij}$), #holders ($m$), #requesters ($n$)

Generic bounds

$$E[T_M] \geq \frac{1}{\mu \lambda \cdot E_p[n]} \cdot E_p \left[ \frac{n}{\bar{g}(n)} \right]$$

$$P\{T_M \leq TTL\} \leq 1 - \frac{1}{E_p[n]} \cdot E_p \left[ n \cdot e^{-\bar{g}(n) \cdot \mu \lambda \cdot TTL} \right]$$
Validation: Performance Prediction

---

**Expected Content Access Delay**

**Content Access Probability**

Network with \( N=10000 \) nodes

*Popularity:* Bounded Pareto with exponent \( a \) and \( 50 \leq n \leq 1000 \)

*Mobility:* \( \lambda_{ij} \) gamma distributed with \( \mu_{\lambda}=1 \) and \( CV_{\lambda} \)

*Availability:* \( g(m,n) \equiv g(n) = 0.2 \cdot n \)
Goal 2: Storage Optimization

- Content A
- Content B
- Content C

Mobile Data Offloading

Not All Content is Created Equal: Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination.

T. Spyropoulos
Goal 2: Storage Optimization

Offloading Algorithm
1. Send each content to some holders (initial placement).
2. Let the holders disseminate the content in an opportunistic / D2D fashion to its requesters

How many holders per content?
• Constraint: Total number of holders
  e.g. total number of transmissions, storage capacity (per node)
• Objectives:

<table>
<thead>
<tr>
<th>Minimize</th>
<th>Maximize</th>
</tr>
</thead>
<tbody>
<tr>
<td>content access delay</td>
<td>content access probability</td>
</tr>
</tbody>
</table>
Goal 2: Storage Optimization

Offloading Algorithm
1. Send each content to some holders (initial placement).
2. Let the holders disseminate the content in an opportunistic / D2D fashion to its requesters

How many holders per content?
- Constraint: Total number of holders
e.g. total number of transmissions, storage capacity (per node)
- Objectives:

\[
\begin{align*}
\min_{\#holders_M} \{ E[T_M] \} & \quad \text{max}_{\#holders_M} \{ P\{T_M \leq TTL\} \} \\
\sum_M \#holders_M = c & \quad \sum_M \#holders_M = c
\end{align*}
\]
Goal 2: Storage Optimization

Offloading Algorithm
1. Send each content to some holders (initial placement).
2. Let the holders disseminate the content in an opportunistic / D2D fashion to its requesters.

How many holders per content?
- Constraint: Total number of holders
  e.g. total number of transmissions, storage capacity (per node)
- Objectives:

\[
\min_{g(n)} \left\{ E_p \left[ \frac{n}{g(n)} \right] \right\}
\]

\[
E_p [g(n)] = c'
\]

Solution:
\[
g(x) = \frac{c}{E_p[\sqrt{n}]} \cdot x
\]

\[
\min_{g(n)} \left\{ E_p \left[ n \cdot e^{-g(n) \cdot \mu \cdot TTL} \right] \right\}
\]

\[
E_p [g(n)] = c'
\]

Convex problem
Validation: Optimization Results

**Expected Content Access Delay**
(Cabspotting trace)

allocation functions: \( g(n) = c_k \cdot n^k \)

x-axis: exponent \( k \)

**Content Access Probability**
(SLAW mobility model)
Extensions

**DIRECTION 1**

**Generic Offloading Model**
- Heterogeneous node types / hierarchy of nodes
- Node cooperation
- Buffer occupancy - Content dropping
- Initial content placement

**DIRECTION 2**

**Detailed Cost Model**
- Different costs for I2D and D2D
- Storing cost
- Delayed content access cost

→ cost optimization
Simple Offloading Model: The Markov Chain

#holders (availability)
#requesters (popularity)
#holders (availability)
#requesters (popularity)

Generic Offloading Model: The Markov Chain

Analysis: Very Complex

cooperation

no cooperation

full buffer
**Generic Offloading Model: Analysis**

*Mean Field* approximation & *Fluid Model* approximations

\[
\frac{dH(t)}{dt} = \rho_c \cdot H(t) \cdot R(t) \cdot \mu \lambda - H(t) \cdot \lambda_d
\]

\[
\frac{dR(t)}{dt} = -H(t) \cdot R(t) \cdot \mu \lambda
\]

Initial conditions: \( H(0) \) and \( R(0) \)

→ **Solutions** for \( H(t) \) and \( R(t) \) (closed form or numerical)

- \( \lambda_d = 0 \)
- \( \lambda_d > 0 \)

\( H(t) \): #holders at time \( t \)
\( R(t) \): #requesters at time \( t \)

**Not All Content is Created Equal:** Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination.
Results: Performance Prediction

\[ P\{T_M \leq TTL\} \text{, Content Access Probability:} \]

\[ 1 - \frac{1}{E_p[n]} \cdot E_p \left[ n \cdot \frac{p_c \cdot n + g(n)}{p_c \cdot n + g(n) \cdot e^{\mu \lambda \cdot (p_c \cdot n + g(n)) \cdot TTL}} \right] \]

\[ E[T_M] \text{, Expected Content Access Delay:} \]

\[ \frac{1}{\mu \lambda \cdot p_c \cdot E_p[n]} \cdot E_p \left[ \ln \left( 1 + \frac{p_c \cdot n}{g(n)} \right) \right] \]
Offloading Algorithm

- Requesters
- Holders

→ Cellular transmission
→ D2D transmission

\( t = 0 \) \( t = \text{TTL} \)
Cost Optimization

Offloading Algorithm
1. The Cellular Network sends each content to some holders (initial placement) at time $t=0$.
2. Let the holders disseminate the content in a D2D fashion to its requesters, till time $t=\text{TTL}$.
3. The Cellular Network sends the content to requesters that have not received it by time $\text{TTL}$.

How many holders per content?

• Costs:
  - $\text{C1}$: initial content placement to a holder (e.g. femtocell, mobile user, etc.)
  - $\text{C2}$: D2D content exchange / storage cost
  - $\text{C3}$: delayed content access cost (i.e. cellular transmission at time $\text{TTL}$)

• Constraint: Total number of holders
  e.g. total number of transmissions at time $t=0$, storage capacity, etc.

• Objectives: minimize total cost
Cost Optimization

Cost of a single content $M$

$$C_M = C_1 \cdot H(0) + C_2 \cdot R(0) \cdot P\{T_M \leq TTL\} + C_3 \cdot R(0) \cdot (1 - P\{T_M \leq TTL\})$$

Optimization Problem

$$\min \sum_{M} C_M \ #holders_M$$

such that

$$\sum_{M} \ #holders_M = c$$

$$\ #holders_M \leq c'$$

Example:

$$g(n) = \frac{1}{\mu \cdot TTL} \cdot \ln \left( \frac{C_3 - C_2}{C_1} \cdot \mu \lambda \cdot TTL \cdot n \right)$$

($p_c=0, \lambda_d=0, c$ and $c'$ large)
THANK YOU !!!
**Definition 5** (Mobility Dependent Allocation). The probability $\pi_{ij}$ a node $i$ to be a holder for a content in which a node $j$ is interested, is related to their contact rate $\lambda_{ij}$ such that $\pi_{ij} = \pi(\lambda_{ij})$, where $\pi(\cdot)$ is a function from $\mathbb{R}^+$ to $[0, 1]$.

**Result:** replace $\mu_{\lambda}$ with

$$
\mu_{\lambda}(\pi) = \frac{E_{\lambda}[\lambda \cdot \pi(\lambda)]}{E_{\lambda}[\pi(\lambda)]}
$$

**Content Access Probability**
Analysis: Markov Chain

\( H \): number of *holders*
\( R \): number of requesters

\[ \lambda_{(m,n) \rightarrow (m+1,n-1)} \approx p_c \cdot H \cdot R \cdot \mu_{\lambda} \]
\[ \lambda_{(m,n) \rightarrow (m,n-1)} \approx (1-p_c) \cdot H \cdot R \cdot \mu_{\lambda} \]
\[ \lambda_{(m,n) \rightarrow (m-1,n)} = H \cdot \lambda_d \]

Mean field approximation
Generic Offloading Model: Results – Single Content

\( H(t) \): number of *holders* at time \( t \)

\( T_d \): content *delivery delay* to a requester

Delivery Probability:

\[
P\{T_d \leq TTL\} = 1 - e^{-\mu \lambda \int_0^{TTL} H(\tau) \, d\tau}
\]

Expected Delivery Delay:

\[
E[T_d] = \int_0^\infty e^{-\mu \lambda \int_0^{t} H(\tau) \, d\tau} \cdot dt
\]

**Examples**

- e.g.1 \( p_c=0, \lambda_d=0 \)

\[
P\{T_d \leq TTL\} = 1 - e^{-\mu \lambda \cdot H_0 \cdot TTL}
\]

- e.g.2 \( \lambda_d=0, R(0)=R_0 \)

\[
E[T_d] = \frac{1}{p_c \cdot R_0 \cdot \mu \lambda} \cdot ln \left( 1 + \frac{p_c \cdot R_0}{H_0} \right)
\]
Generic Offloading Model: Results – Multiple Contents

Communication Scenario - Initial Allocation:
- For each content with popularity $n$, the CP assigns $H_0 = g(n)$ initial holders
- $g(n): [1,N] \rightarrow [0,N]

Delivery Probability:
$$P\{T_d \leq TTL\} = 1 - \frac{1}{E_p[n]} \cdot E_p\left[n \cdot e^{-\mu \cdot \int_0^{TTL} H(\tau|n) d\tau}\right]$$

Expected Delivery Delay:
$$E[T_d] = \frac{1}{E_p[n]} \cdot E_p\left[n \cdot \int_0^{\infty} e^{-\mu \cdot \int_0^{t} H(\tau|n) d\tau}\right]$$

Examples

e.g.1 $p_c = 0, \lambda_d = 0$
$$P\{T_d \leq TTL\} = 1 - \frac{1}{E_p[n]} \cdot E_p\left[n \cdot e^{-\mu \cdot g(n) \cdot TTL}\right]$$

e.g.2 $\lambda_d = 0, R(0) = R_0$
$$E[T_d] = \frac{1}{\mu \lambda \cdot E_p[n]} \cdot E_p\left[n \cdot g(n)\right]$$
Generic Offloading Model: Results - Tables

<table>
<thead>
<tr>
<th>Cooperation</th>
<th>$H(t)$</th>
<th>$P{T_d \leq TTL}$</th>
<th>$E[T_d]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic scenario</td>
<td>$(p_c \cdot R_s + H_0) = \frac{p_c \cdot R_s \cdot (p_c \cdot R_s + H_0)}{p_c \cdot R_s + H_0 \cdot e^{\mu \lambda \cdot (p_c \cdot R_s + H_0) \cdot t}}$</td>
<td>$1 - \frac{p_c \cdot R_s + H_0}{p_c \cdot R_s + H_0 \cdot e^{\mu \lambda \cdot (p_c \cdot R_s + H_0) \cdot TTL}}$</td>
<td>$\frac{1}{\mu \lambda \cdot p_c \cdot R_s \cdot H_0} \cdot \ln \left(1 + \frac{p_c \cdot R_s}{H_0}\right)$</td>
</tr>
<tr>
<td>Base scenario</td>
<td>$H_0$</td>
<td>$1 - e^{-\mu \lambda \cdot H_0 \cdot TTL}$</td>
<td>$\frac{1}{\mu \lambda \cdot H_0}$</td>
</tr>
</tbody>
</table>

Table 1: Solution for the number of holders over time $H(t)$ (Lemma 3.1), the content delivery probability (Result 1) and expected delivery delay (Result 2) for scenarios with $\lambda_d = 0$.

<table>
<thead>
<tr>
<th>Cooperation</th>
<th>$P{T_d \leq TTL}$</th>
<th>$E[T_d]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic scenario</td>
<td>$1 - \frac{1}{E_p[n]} \cdot E_p \left[ \frac{p_c \cdot n + g(n)}{p_c \cdot n + g(n) \cdot e^{\mu \lambda \cdot (p_c \cdot n + g(n)) \cdot TTL}} \right]$</td>
<td>$\frac{1}{\mu \lambda \cdot p_c \cdot E_p[n]} \cdot E_p \left[ \ln \left(1 + \frac{p_c \cdot n}{g(n)}\right) \right]$</td>
</tr>
<tr>
<td>Base scenario</td>
<td>$1 - \frac{1}{E_p[n]} \cdot E_p \left[ n \cdot e^{-\mu \lambda \cdot g(n) \cdot TTL} \right]$</td>
<td>$\frac{1}{\mu \lambda \cdot E_p[n]} \cdot E_p \left[ \frac{n}{g(n)} \right]$</td>
</tr>
</tbody>
</table>

Table 2: Solution for the content delivery probability and expected delivery delay of a random request (Result 4) for scenarios with $\lambda_d = 0$. 

Not All Content is Created Equal: Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination.
Simulation Results: Single Content

Not All Content is Created Equal: Effect of Popularity and Availability for Opportunistic Device-based Storage and Dissemination.

T. Spyropoulos
Cost Optimization:
Solution for *example case* with no total holders constraint

\[ g^*(n) = \begin{cases} 
N_{IN} & , n \geq R_{(\text{max})} \\
\frac{1}{\gamma} \cdot \ln \left( \frac{\Phi_3 - \Phi_2}{\Phi_1} \cdot \gamma \cdot n \right) & , R_{(\text{min})} \leq n < R_{(\text{max})} \\
0 & , n < R_{(\text{min})}
\end{cases} \]

\[ \gamma = \mu \lambda \cdot TTL \]

\[ R_{(\text{min})} = \frac{1}{\gamma} \cdot \frac{\Phi_1}{\Phi_3 - \Phi_2} \]

\[ R_{(\text{max})} = \frac{1}{\gamma} \cdot \frac{\Phi_1}{\Phi_3 - \Phi_2} \cdot e^{\gamma \cdot N_{IN}} \]
Validation: Cost Optimization

\[ C_1 = C_3 = 50 \ C_2 \]

Population = \( R(0) = 100 \)

\[ \gamma = \mu_\lambda \cdot TTL \]