Temporal Coverage Based Content Distribution in Heterogeneous Smart Device Networks

Wei Peng   Feng Li   Xukai Zou

Indiana University-Purdue University Indianapolis (IUPUI.edu)

Content Distribution in Heterogeneous Smart Device Network

heterogeneous smart device network

- partial cellular coverage
  - opportunistic proximate channel available on all devices
  - only *some* has persistent cellular links
- why
  - users disable cellular links for cost concerns
  - some tablets do not have cellular capability

- scenario
  - content injected into network through devices with cellular channel (the "seeds")
  - propagate through proximate channel when devices come close to each other
  - no central coordination due to lack of cellular channel for some devices
- objective
  - all proximately reachable devices be covered
  - reduce cost: number of proximate channel copies

- eager multiple forwarding/flooding
  - forward over proximate channel upon encounter
  - delivery delay is minimized, but cost can be high
- eager k forwarding
  - forward for the first k encounters
  - how to select k? coverage?
- random forwarding
  - forward by flipping coins
  - how to select the odds?

Content Distribution based on Temporal Covering Set

Ideas

- restrict forwarding:
  - not by k
  - but by membership in a **temporal covering set**
- temporal covering set: a subset of devices with
  - strong internal connectivity
    - so that content can be propagated
  - full external coverage
    - so that every node can receive the
- strength of connection $\Rightarrow$ temporal quality of proximate channel

Proximate Channel Temporal Quality

Previous: Average Inter-encounter Interval

- given u's past encounters $[s_1^{u,v}, e_1^{u,v}], \ldots, [s_{k_{u,v}}^{u,v}, e_{k_{u,v}}^{u,v}]$ with v
- u estimates the **temporal quality** of its proximate channel with v
  - in terms of the proximate channel's potential of forwarding the content timely
- a straightforward metric: average inter-encounter interval

$$\frac{1}{k_{u,v}-1} \sum_{i=1}^{k_{u,v}-1} \left(s_{i+1}^{u,v} - e_i^{u,v}\right)$$

  - disadvantage: not capturing uncertainty of proximate channel quality

Proximate Channel Temporal Quality

Uncertainty about Estimation: An Example

- 10 groups of inter-encounter interval records
  - Group i ($i \in \{1, 2, \ldots, 10\}$) consists of 2^i pairs of interleaved 100 and 200 (units of time)
  - Example: Group 2 is "100, 200, 100, 200"
- the desired quality: 110
- average inter-encounter interval is 150 for all groups
- however, by intuition:
  - periodically has 100 that satisfies desired quality
  - Group 10 has more certainty than Group 1

Proximate Channel Temporal Quality

Solution: T-coverage Quality $d_u^T(v)$

- idea: using KDE (kernel density estimation) of u's inter-encounter intervals with v
- smoothing kernel function $\hat{f}_{u,v}(x)$
  - with Epanechnikov kernel $K(x) = \frac{3}{4}(1-x^2)\mathbf{1}_{|x|\leq 1}$

$$\hat{f}_{u,v}(x) = \frac{1}{k_{u,v}-1} \sum_{i=1}^{k_{u,v}-1} K(x - (s_{i+1}^{u,v} - e_i^{u,v}))$$

- T-coverage quality metric $d_u^T(v)$

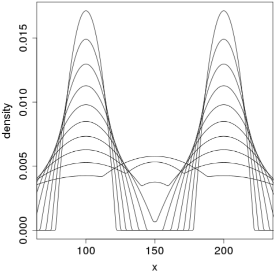
$$d_u^T(v) = \int_{-\infty}^{T} \hat{f}_{u,v}(x)\mathrm{d}x$$

  - T: a time-domain quality threshold to filter out sporadic or long-delay opportunistic links between nodes
    - without T, $d_u^T(v)$ always integrates to 1 $\Rightarrow$ not usable
    - greater $d_u^T(v)$ $\Rightarrow$ better chance for timely content forwarding/delivery

Proximate Channel Temporal Quality

Back to the Example

T-coverage quality metric $d_u^T(v)$ with $T = 110$



i in 2^i	$d_u^T(v)$	i in 2^i	$d_u^T(v)$
1	0.293	6	0.346
2	0.303	7	0.360
3	0.312	8	0.375
4	0.323	9	0.391
5	0.334	10	0.410

Temporally Covering Set

- U: all devices; U_c: seeds
- $\mathcal{D}_T \subset U$: temporally covering set with temporal threshold of T (T-covering set)
  - (Coverage) For each node $u \in U$, either $u \in \mathcal{D}_T$ or there is a node $v \in \mathcal{D}_T$ such that u is T-covered by v.
  - (Connectivity) For each node $u \in \mathcal{D}_T$, either u is a seed (i.e., $u \in U_c$), or there is a seed $v \in U_c$ (i.e., v is equipped with cellular data channel) such that there is a path (i.e., a chain of consecutively T-covered nodes) from v to u.
- T-dominators and T-dominatees
- by Connectivity, non-seed dominators are also dominatees

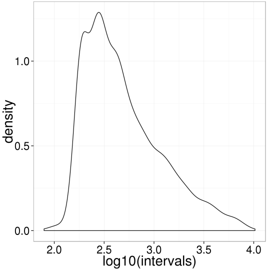
Dominator Election Algorithm

- each u locally keeps:
  - downstream list $L_{\downarrow}(u)$: nodes that are best dominated by u
  - upstream list $L_{\uparrow}(u)$: u's upstream to some seed
- information exchange when u and v meet
  - u to v
    - u sends its seed/dominator status to v
    - u sends $L_{\uparrow}(u)$ and $L_{\downarrow}(u)$ to v
    - u receives $\{d_v^T(w) | w \in L_{\downarrow}(u)\}$ and $\{d_v^T(w) | w \in L_{\uparrow}(u)\}$ from v
  - similarly for v to u
- u locally adjusts its dominator/non-dominator status algorithm detail
  - update $L_{\uparrow}(u)$ and $L_{\downarrow}(u)$
  - turn dominator if both $L_{\uparrow}(u)$ and $L_{\downarrow}(u)$ nonempty
- T-covering set emerges out of the collective effect of such local status adjustment

Bluetooth encounter dataset sigcomm2009

- from CRAWDAD wireless dataset archive
- timestamped periodic Bluetooth proximity device discovery records of 48 regularly meeting nodes
- 2, 4, 8 seeds

smoothed density distribution for inter-encounter intervals



take quality threshold $T = 1,000$

- emulti: eager multiple forwarding; as baseline
- esingle: eager k-forwarding with $k = 1$
- random50: random forwarding with 50% forwarding probability
- tdom: T-coverage-based forwarding ($T = 1,000$)
- tdom50: T-coverage-based forwarding ($T = 1,000$) with 50% forwarding probability

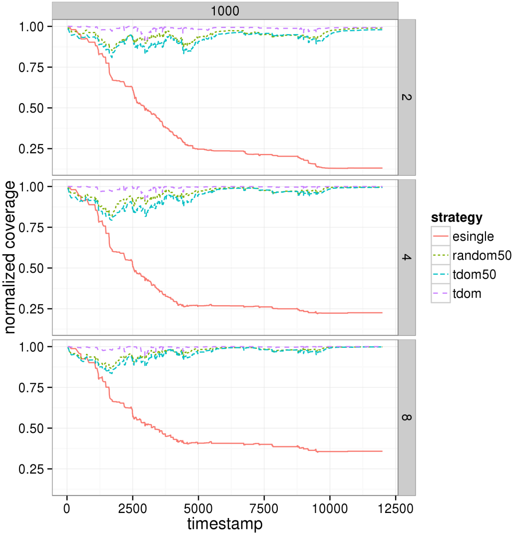
average content delivery delay comparing with emulti

	esingle	random50	tdom50	tdom
2	5577	271	397	81
4	5530	199	306	29
8	4725	173	241	25

tdom has small delay

Evaluation

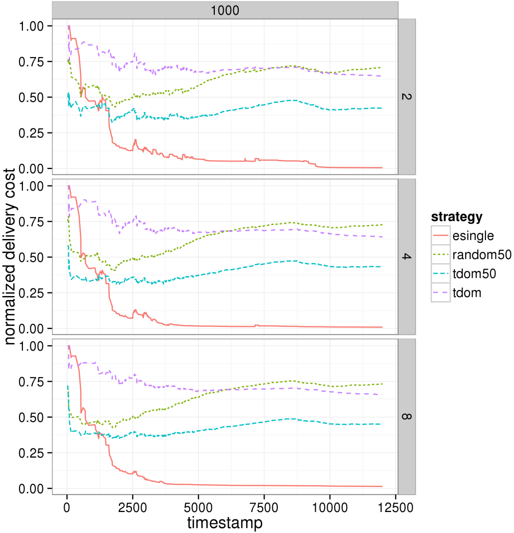
Result: Coverage (Normalized by emulti)



tdom has coverage (almost) the same with emulti. . .

Evaluation

Result: Cost (Normalized by emulti)



...at the cost of about 75%

Summary

- KDE-based temporal quality metric captures uncertainty with a single number
- comparing with flooding, localized dominator election algorithm has complete coverage at a lower cost

Q&A

Thank You

Backup Slides

Dominator Election Algorithm ▸ back

u adjusts its dominator status after encountering v

1: ▶ only consider quality opportunistic links
2: **if** $d_u^T(v) > 0$ **then**
3: ▶ u updates $L_\uparrow(u)$ and $L_\downarrow(u)$
4: **if** v is a dominator **then**
5: **if** $v \in U_c$ or u is a non-dominator **then**
6: $L_\uparrow(u) \leftarrow L_\uparrow(u) \cup \{v\}$
7: $L_\downarrow(u) \leftarrow L_\downarrow(u) \setminus \{v\}$
8: **end if**
9: **for** $w \in L_\downarrow(u)$ **do**
10: **if** $d_v^T(w) > d_u^T(w)$ **then** ▶ if w is better T-dominated by v than by u
11: $L_\downarrow(u) \leftarrow L_\downarrow(u) \setminus \{w\}$
12: **end if**
13: **end for**
14: **else if** v is a non-dominator **then**
15: $L_\uparrow(u) \leftarrow L_\uparrow(u) \setminus \{v\}$
16: $x \leftarrow$ TRUE ▶ $x =$ TRUE if u is v's best T-dominator
17: **for** $w \in L_\uparrow(u)$ **do**
18: **if** $d_v^T(w) > d_u^T(v)$ **then** ▶ if v is better T-dominated by w than by u
19: $x \leftarrow$ FALSE
20: go to 23
21: **end if**
22: **end for**
23: **if** $x =$ TRUE **then** ▶ u is v's best dominator
24: $L_\downarrow(u) \leftarrow L_\downarrow(u) \cup \{v\}$
25: **end if**
26: **end if**
27: ▶ u sets its dominator status based on whether $L_\downarrow(u)$ and $L_\uparrow(u)$ are empty
28: **if** $u \notin U_c$ **then** ▶ only non-seeds change dominator status
29: **if** $L_\downarrow(u) = \emptyset$ **then**
30: u turns a non-dominator
31: **else if** $L_\uparrow(u) \neq \emptyset$ and $L_\downarrow(u) \neq \emptyset$ **then**
32: u turns a dominator
33: **end if**
34: **end if**
35: **end if**