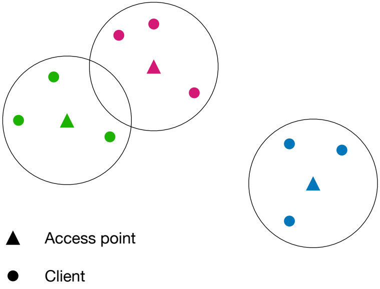
Passive learning of the interference graph of a wireless network

Stark Draper
University of Toronto

work with Jing Yang and Robert Nowak

BIRS
March 2024

Wireless local access network, e.g., 802.11



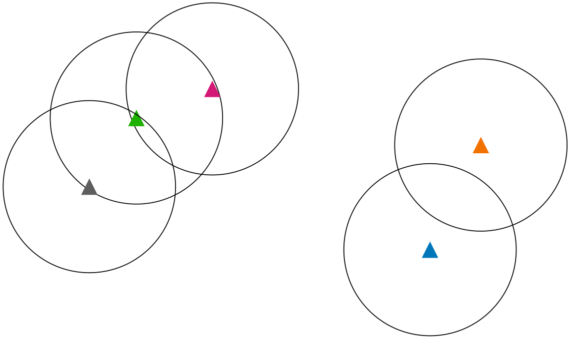
- Knowledge of interference among access points (APs), useful to set network configuration: dynamic channel assignments, transmit power control, scheduling, etc..
- Interference changes over time as environment moves

Estimating interference among APs

- Motivated by work in CS systems. Shrivastava et al. "PIE in the sky: Online passive interference estimation for enterprise WLANs," *USENIX Conf. Net. Sys. Design Implementation*.
  - Earlier work injected traffic into network (active probing) to characterize inference.
  - Shrivastava leveraged CSMA/CA and Ack/Nak protocols to characterize interference
- In paper, small test-bed studies only. Curious about how scaled up to large networks.
- Some analogies with a type of group testing problem.
- Main contribution: Quantify the required observation time as a function of network size and topological connectivity.
  - Formulated PIE as a conflict graph learning problem.
  - Edge set represents pairwise interference among APs.
  - Recover the interference graph with as few measurements as possible.

Direct interference 1

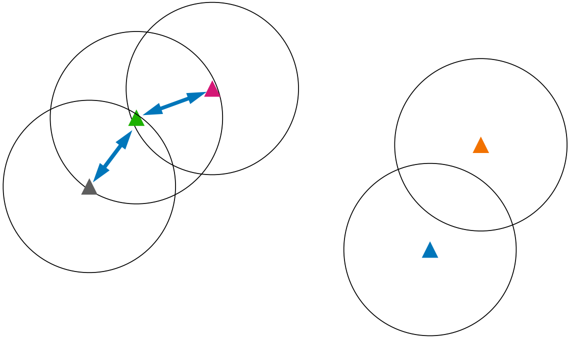
- CSMA/CA: AP holds its transmission if channel is busy.



- APs w/in carrier sensing range: Never transmit at same time.

Direct interference 2

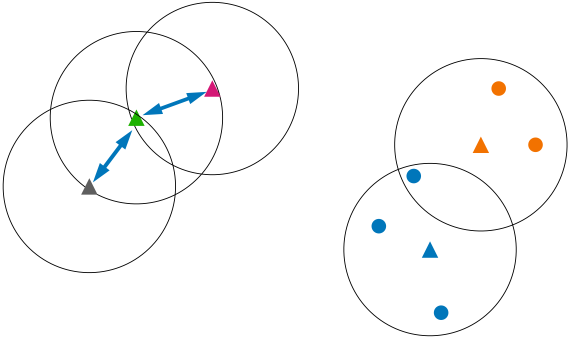
- CSMA/CA: AP holds its transmission if channel is busy.



- APs w/in carrier sensing range: Never transmit at same time.
- We model carrier sensing as reciprocal (undirected graph).

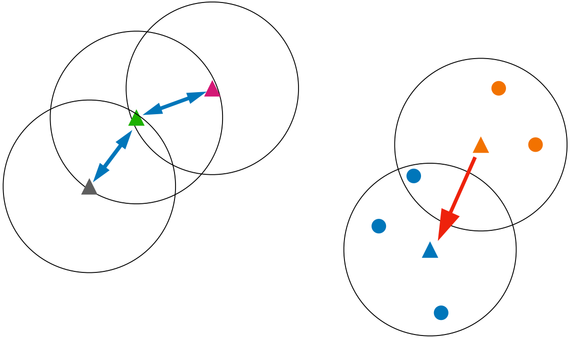
Hidden interference 1: Hidden terminal problem

- Hidden: Out of carrier sensing range / shielded by obstruction



Hidden interference 2: Hidden terminal problem

- Hidden: Out of carrier sensing range / shielded by obstruction



- Collisions may happen, depends on receiver's location
- Interference is asymmetric (directed graph)
- Collisions can be detected through Ack/Nak mechanism

Interference graphs: Direct and hidden

- Use graph $G = (\mathcal{V}, \mathcal{E}_D, \mathcal{E}_H)$ to represent the interference where $\mathcal{V}$, $|\mathcal{V}| = n$ is the set of access points.
- $G_D = (\mathcal{V}, \mathcal{E}_D)$: Graph of direct interference among APs.
  - Carrier sensed since direct.
  - Data: Network activation pattern $X \in \{0, 1\}^n$, 1-active, 0-inactive
  - Undirected graph (reciprocal sensing)
- $G_H = (\mathcal{V}, \mathcal{E}_H)$: hidden interference to another AP's clients
  - When APs cannot hear each other, no direct path, e.g., building in way.
  - AP may interfere only with a subset of another AP's clients
  - Data: activation pattern X; feedback information $Y \in \{0, 1\}^n$ where 1-Ack, 0-Nak.
  - Directed graph (asymmetric interference)
- Objective: given k observations of X and Y, recover G_D, G_H.

System model

- n access points.
- Synchronized time-slotted system.
- In each slot, with probability p, an AP has data to send (i.i.d. across slots)
- $\mathcal{N}_i$: set of direct interferers for AP i, $|\mathcal{N}_i| \leq d$
- Uniform contention for channels among APs.
- $\mathcal{S}_i$: set of hidden interferers for AP i, $|\mathcal{S}_i| < s$
- Probability that a client associated with AP i is interfered with by AP j, $j \in \mathcal{S}_i$, is $p_{ji} \geq p_{\min}$
- Consider static channel states.

Inferring direct interference G_D

Algorithm:

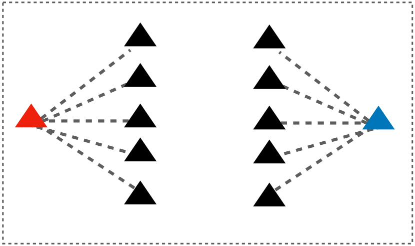
- If concurrent Tx from both i and j then $(i, j) \notin \mathcal{E}_D$
- Start from fully connected graph, gradually remove edges

Inferring direct interference G_D

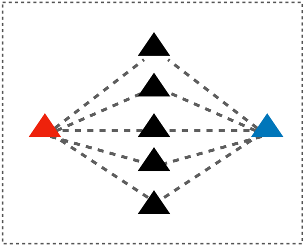
Algorithm:

- If concurrent Tx from both i and j then $(i, j) \notin \mathcal{E}_D$
- Start from fully connected graph, gradually remove edges

Aspect of the analysis is that activation patterns not i.i.d.. across network, the conflict graph introduces dependencies.



G_{D0}



G_{D1}

Analysis and results

- Lemma: In k steps $\Pr(\hat{G}_D \neq G_D) \leq \binom{n}{2}\left(1 - \frac{p^2}{(d+1)^2}\right)^k$
- Analysis deals with dependence in activation patterns:
  - Bound # edges by $\binom{n}{2}$
  - Pr a non-interfering pair does not simultaneously transmit in a given slot $\leq (1 - p^2/(d+1)^2)$
- Theorem 1:
  - $\Pr(\hat{G}_D \neq G_D) < \delta$ if $k \geq \frac{1}{\log \frac{1}{1-p^2/(d+1)^2}}\left(\log \binom{n}{2} + \log \frac{1}{\delta}\right)$.
- $k = O(d^2 \log n)$ when $p^2/d^2 \to 0$.

Minimax lower bound

Theorem 2: For any α, $0 < \alpha < 1/8$ and under some mild conditions on d and n, if

$$k \leq \frac{\alpha d^2}{2 + \frac{1}{1-p}} \log n$$

then

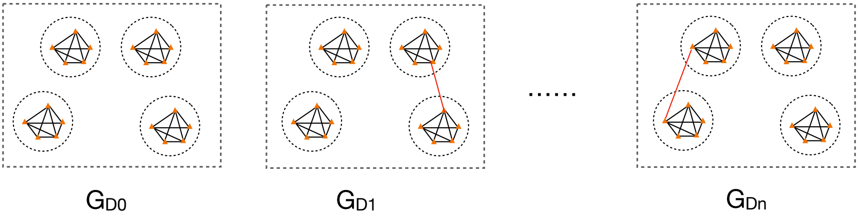
$$\min_{\hat{G}_D \in \mathcal{G}_D} \max_{G_D \in \mathcal{G}_D} \Pr(\hat{G}_D \neq G_D) > 0$$

.

- So if k scale slower then $d^2 \log n$ will have an error.
- Upper bound matches lower bound up to a constant.

Proof idea: Develop a simple family of networks

1. Pick simple, but hard-to-distinguish among, set of networks.
2. Turn into hypothesis testing problem where there are $(n+1)$ graphs $G_{D0}, G_{D1}, \ldots G_{Dn}$.
3. Characterize distribution of activation patterns $P_i(X)$ for all $i \in \{0, 1, \ldots n\}$.
4. $D_{\mathrm{KL}}(P_0 \| P_i) = D_{\mathrm{KL}}(P_0 \| P_1)$ for all $i \in [n]$.
5. Adapt bound from Tsybakov '08 to lower bound hypothesis test; also lower bounds original (more difficult) problem.



Result: Adapted from Tsybakov, '08, Thm. 2.5

Theorem: Let $k \in \mathbb{Z}^+$, $M \geq 2$, $\{G_{D0}, \ldots, G_{DM}\} \in \mathcal{G}_D$ be such that

1. $D_L(G_{Di}, G_{Dj}) \geq 2r$, for $0 \leq i < j \leq M$, where D_L is the Levenshtein ("edit") distance.
2. $\frac{k}{M} \sum_{i=1}^{M} D_{KL}(P_i \| P_0) \leq \alpha \log M$, with $0 < \alpha < 1/8$.

Then,

$$\begin{aligned}\inf_{\hat{G}_D \in \mathcal{G}_D} \sup_{G_D \in \mathcal{G}_D} \Pr[D_L(\hat{G}_D, G_D) \geq r; G_D] &\geq \inf_{\hat{G}_D \in \mathcal{G}_D} \max_i \Pr[D_L(\hat{G}_D, G_{Di}) \geq r; G_{Di}] \\ &\geq \frac{\sqrt{M}}{1+\sqrt{M}} \left(1 - 2\alpha - \sqrt{\frac{2\alpha}{\log M}}\right) > 0.\end{aligned}$$

Inferring hidden interference G_H: Example

- Consider AP j. At time t if the feedback $Y_j^t = 0$ then the transmission was unsuccessful.
- In example let set of hidden interferers be $\mathcal{S}_j = \{i_1, i_2\}$

Feedback	APs active in slot
$Y_j^{(1)} = 0$	$\{1, 2, 7, i_i, 9\}$
$Y_j^{(2)} = 1$	(ignore)
$Y_j^{(3)} = 0$	$\{1, 3, 5, i_2, 10\}$
$Y_j^{(4)} = 0$	$\{3, 5, i_i, 8\}$

Note that $\{i_i, i_2\}$ and $\{1, 3\}$ both intersect all failures (and $(2, 5)$ does too). Can't yet determine $\mathcal{S}_j$. So, wait...

Inferring hidden interference G_H: Example

- Consider AP j. At time t if the feedback $Y_j^t = 0$ then the transmission was unsuccessful.
- In example let set of hidden interferers be $\mathcal{S}_j = \{i_1, i_2\}$

Feedback	APs active in slot
$Y_j^{(1)} = 0$	$\{1, 2, 7, i_i, 9\}$
$Y_j^{(2)} = 1$	(ignore)
$Y_j^{(3)} = 0$	$\{1, 3, 5, i_2, 10\}$
$Y_j^{(4)} = 0$	$\{3, 5, i_i, 8\}$

Note that $\{i_i, i_2\}$ and $\{1, 3\}$ both intersect all failures (and $(2, 5)$ does too). Can't yet determine $\mathcal{S}_j$. So, wait...

Feedback	APs active in slot
$Y_j^{(5)} = 0$	$\{2, i_2, 11\}$

Now the set $\mathcal{S}_j = \{i_1, i_2\}$ is the unique "minimum hitting set"

Minimum hitting set based approach

Def: Given a collection of subsets, a set which intersects all subsets in at least one element is called a hitting set; a minimum hitting set is a hitting set of smallest size.

Minimum hitting set based approach

Def: Given a collection of subsets, a set which intersects all subsets in at least one element is called a hitting set; a minimum hitting set is a hitting set of smallest size.

1. Given k observations define $\mathcal{K}_j(k)$ to be the slots in which AP j's transmissions failed, i.e.,

$$\mathcal{K}_j(k) = \{t \in \{1, 2, \ldots, k\} | Y_j(t) = 0\}$$

Minimum hitting set based approach

Def: Given a collection of subsets, a set which intersects all subsets in at least one element is called a hitting set; a minimum hitting set is a hitting set of smallest size.

1. Given k observations define $\mathcal{K}_j(k)$ to be the slots in which AP j's transmissions failed, i.e.,

$$\mathcal{K}_j(k) = \{t \in \{1, 2, \ldots, k\} | Y_j(t) = 0\}$$

2. For each such failure ($t \in \mathcal{K}_j(k)$) define $\mathcal{S}_j^t$ to be the set of *candidate* hidden interferers, i.e.,.

$$\mathcal{S}_j^t = \{i \in \mathcal{V} | i \neq j, X_i(t) = 1\}$$

Minimum hitting set based approach

Def: Given a collection of subsets, a set which intersects all subsets in at least one element is called a hitting set; a minimum hitting set is a hitting set of smallest size.

1. Given k observations define $\mathcal{K}_j(k)$ to be the slots in which AP j's transmissions failed, i.e.,

$$\mathcal{K}_j(k) = \{t \in \{1, 2, \ldots, k\} | Y_j(t) = 0\}$$

2. For each such failure ($t \in \mathcal{K}_j(k)$) define $\mathcal{S}_j^t$ to be the set of *candidate* hidden interferers, i.e.,.

$$\mathcal{S}_j^t = \{i \in \mathcal{V} | i \neq j, X_i(t) = 1\}$$

3. Our estimate of $\mathcal{S}_j$ is the minimum hitting set

$$\hat{\mathcal{S}}_j(k) = \arg\min_{\mathcal{S} \subseteq \mathcal{V}} \left\{ |\mathcal{S}| \middle| \mathcal{S} \cap \mathcal{S}_j^t \neq \emptyset \ \forall \ t \in \mathcal{K}_j(k) \right\}$$

Analysis

Lemma: At any k for any minimal hitting set $\hat{\mathcal{S}}_j(k)$ and any $i \in \mathcal{S}_j$,

$$\Pr[i \notin \hat{\mathcal{S}}_j(k)] \leq \left(1 - \frac{p^2}{(d+1)^2} p_{ij} (1-p)^s\right)^k$$

- $p^2/(d+1)^2$ same as earlier – lower bound on APs i, j both active and therefore *may* interfere
- p_{ij} probability AP i interferes with AP j transmission
- $(1-p)^s$ probability that any other candidate minimal hitting set of hidden interferes have nothing to send in a particular round (and so cannot interfere)

Apply union bound to get

$$\begin{aligned} Pr[\hat{G}_H \neq G_H] &= \Pr[\cup_{j \in [n]} \cup_{i \in \mathcal{S}_j} i \notin \hat{\mathcal{S}}_j(k)] \\ &\leq ns \left(1 - \frac{p^2}{(d+1)^2} p_{\min} (1-p)^s\right)^k \end{aligned}$$

Results

- Theorem 3: Let $\delta > 0$ and let

$$k \geq \frac{1}{-\log\left(1 - \frac{p^2}{(d+1)^2} p_{\min}(1-p)^s\right)} \left(\log(ns) + \log\frac{1}{\delta}\right)$$

then $\Pr[\hat{G}_H = G_H] > 1 - \delta$.

- As d gets large then simplifies to

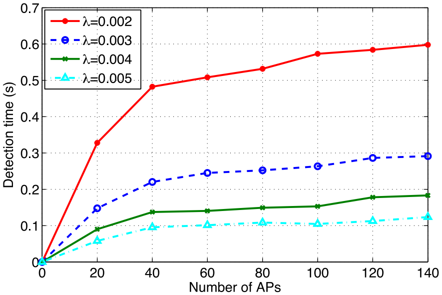
$$k = O\left(\frac{d^2}{p^2(1-p)^s p_{\min}} \log n\right)$$

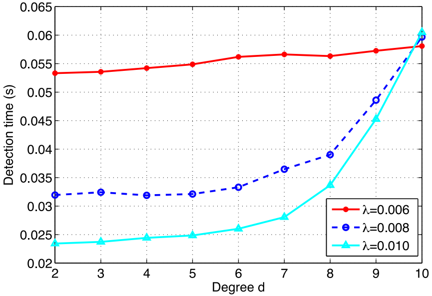
- Min-Max lower bound (Thm 4): We also derive a min-max lower bound that show that, as d gets large, if k scales slower than $\left(\frac{d^2}{(1-p)^{s-1} p_{\min}} \log n\right)$ the probability of error will be bounded away from zero.

Numerical results: Direct interferers

More realistic than assumptions made in analysis:

- Simulated wireless, path loss and shadowing.
- Poisson packet arrival process, enqueued at APs, not i.i.d.
- Grid of $50m \times 50m$ cells, client at center, AP placed randomly
- Here $d = 6$ direct interferers on left, and 4×15 grid on right



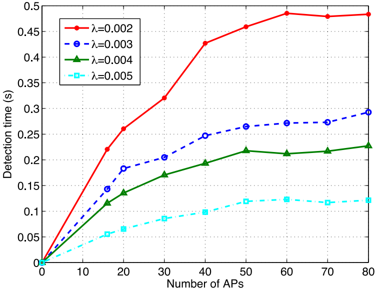


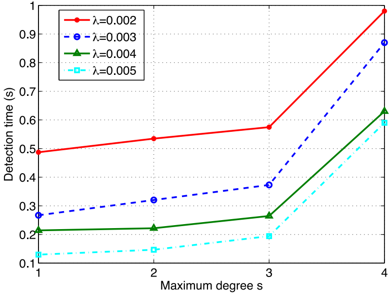
Scaling on left, in # APs (n), sub-linear (logarithmic).

Scaling on right, in max # interferers, super-linear (quadratic)

Numerical results: Hidden interferers

- Same set-up as previous slide
- $s = 1$ hidden interferers on left
- 4×15 grid on right, variable s





As on previous slide, sub-linear scaling in # APs, and super-linear in # (hidden) interferers

Conclusions

Recap

- Formulated the "passive interference estimation" approach as a statistical learning problem
- Two sub-problems: direct and indirect (hidden) interference
- Developed practical algorithms and proved optimality.

Comments

- Can be thought of as a type of group testing problem.. The interferers (direct or hidden) are the "defects" to be determined. There are multiple group testing problems in parallel, i.e.., the set of interferers for each AP. We don't control the test vectors (the randomly generated activation / feedback patterns).
- Scaling in hidden interference, $O(d^2 \log n / p^2 p_{\min} (1-p)^s)$: has a poor constant for high traffic networks (p close to one).
- We developed results for *exact* graph recovery, inexact may often suffice