Sensor Selection

CSE590
Sensor Selection

- Which sensor to turn on, to best sense the event of interest?

- Simple setting: target detection
  - Turn on the one closest to the target
  - Often triggered by the target, etc, acoustic, seismic.

- Not so easy setting:
  - Sense the temperature distribution
  - Large size spatial events (fire)
  - detection large network failures, spatial cuts.
Problem I: Detect significant events

- "large" events: affect $\varepsilon$ fraction of the network
- Goal: select a small $O(1)$ set of monitoring sensors to detect ("catch") these large events.
- First consider one big event and random sampling.
Random sampling

• Given a big event $R \subseteq U$, $|R| > \varepsilon |U|$.
• Take a random sample of $m = \left( \frac{1}{\varepsilon} \ln \frac{1}{\delta} \right)$ sensors.
• Claim: one of them is in $R$ with prob $1-\delta$.
• Proof:

  $\text{Prob\{no sample in } R\}$
  
  $\leq (1 - \varepsilon) \frac{1}{\varepsilon} \ln \frac{1}{\delta} \leq e^{-\ln \frac{1}{\delta}} = \delta.$

Note: the sample is specific for $R$. 
Detect significant events

- How about a universal sample that can detect all possible large events?
- The caveat: we can only hope to catch nicely shaped events.
- Suppose we activate $m$ sensors, but the event includes the rest of the sensors. $\Rightarrow m$ has to be $\Omega(n)$. 

![Diagram showing a universal sample and the condition for $m$.]
Complexity of Shape: VC-dimension

• A circle is simpler than a rectangle, why?
• Range space \((X, R)\)
  – \(X\) is a ground set
  – \(R\) is a set of subsets (ranges) of \(X\)
• \(A \subseteq X\) is shattered by \(R\) if all possible subsets of \(A\) can be obtained by intersecting \(A\) with an \(r \in R\)
• The **VC-Dimension of \((X, R)\)** is the cardinality of the largest set \(A\) that can be shattered by \(R\)
Examples of VC-dimension

• \((X, R)\), \(X=\) a set of points in \(\mathbb{R}^1\),
• All intervals on the line: 2
• Half spaces: 3

\[
\begin{align*}
&\text{\includegraphics[width=\textwidth]{example图.jpg}}
\end{align*}
\]
Examples of VC-dimension

- $(X, R)$, $X$=a set of points in $\mathbb{R}^2$,
- Circles (of all possible radius and center): 3
Examples of VC-dimension

• $(X, R)$, $X$=a set of points in $\mathbb{R}^2$,
• Axis-aligned rectangles: 4
• Convex polygon: infinite, why?
Random Sampling

- A subset $B$ is an $\varepsilon$-net if it has a non-empty intersection with any range of size $\geq \varepsilon n$
- The $\varepsilon$-net theorem: A set of $m$ independent random samples is an $\varepsilon$-net with prob $1-\delta$, if
  \[ m \geq \max \left( \frac{8d}{\varepsilon} \log \frac{8d}{\varepsilon}, \frac{4}{\varepsilon} \log \frac{2}{\delta} \right) \]
- Conclusion: $O(1)$ random samples can be used to detect large events with prob $1-\delta$.
- Problem solved?
Problem with random samples

• **False alarm**: no guarantee that what we detect is a large event. Need if and only if.

• **No size estimation**: how large is the event?

• Select more samples?

• **ε-sample**: \[ \Theta \left( \frac{d}{\varepsilon^2} \log \frac{d}{\varepsilon} + \frac{1}{\varepsilon^2} \log \frac{1}{\delta} \right) \]

• One can estimate the size of any range within additive error \( O(\varepsilon n) \).

• But it is much larger, \( O(1/\varepsilon^2) \) v.s. \( O(1/\varepsilon) \)
Symmetric Difference

- \( D_1 \oplus D_2 := (D_1 \cup D_2) \setminus (D_1 \cap D_2) \)
- Suppose \((X, R)\) has VC dimension \(d\)
- \(R' := \{r_1 \oplus r_2 \mid r_1, r_2 \in R\}\)
- Then \((X, R')\) has VC dimension \(d' := O(d \log d)\)
Algorithm with two-sided guarantee

• Suppose the events are all disk shape.
• Construct an $\varepsilon/4$-net $M$ with respect to the symmetric difference ranges of dimension $d'$.
• Say $T \subseteq M$ detect 1.
• Compute a disk including $T$ and excluding $M \setminus T$.
• The disk contains total $K$ sensors.
• If $K \geq 3\varepsilon n/4$, then report the event as large event, with $K$ as its predicted size. Otherwise, the event is not a large event.
Proof

• Suppose $D^*$ is the real event and $|D^*| = K^*$.
• $T \subseteq D \cap D^*$, thus $M \cap (D \oplus D^*) = \phi$
• By $\varepsilon$-net property, $|D \oplus D^*| \leq \varepsilon n/4$
• $|K^* - K| \leq |D \oplus D^*| \leq \varepsilon n/4$ #

• All events of size $\geq \varepsilon n$ are caught.
• No events of size $\leq \varepsilon n/2$ are reported.
General algorithm

• What if the events are not disks?
• Note that the algorithm only requires set-theoretical properties, not the shape of the events.
• D can be replaced by any geometric shape of VC dimension at most d
• In practice, calculate simply shaped regions (rectangles, ellipses) that separate T from M\T
Simulations
Simulations
Simulations
Problem II: Detect Network Cuts

- Network partition by a linear cut.
- **ε-cut**: network partition where ε-fraction of sensors cut off from the base station.
- **Linear ε-cut**: partition defined by a line.
- Two equivalent views: physical disabling of sensors, or communication disruption along the cut.
Resource Constraints

- Continuously polling each and every sensor is infeasible.
- Communication and bandwidth are both limited.
- Need for a low overhead, low latency, high confidence, cut detection scheme.
A Model of Monitoring

• The base station knows the locations of all sensors.
• A small subset of nodes is designated as sentinels.
• Each live sentinel node sends a heartbeat message to the base station per epoch.
• No message from the dead sentinels.
• The live/dead signature vector at the base station is sufficient to decide if an $\varepsilon$-cut has occurred.
Sampling Methods

• By the $\varepsilon$-net theory, an $O(1/\varepsilon \log 1/\delta)$ size random sample of nodes can act as a detection set.

• But this scheme has a large rate of false positives.

• $\varepsilon$-approximation sample eliminates the false positives, but it requires too large a sentinel set.

• For instance, with $\varepsilon = 0.1$ and $\delta = 0.05$, the $\varepsilon$-approximation has size is at least 10,000.
\( \varepsilon \)-cuts and approximation guarantees

- Cuts must be defined as a fraction of the network size.
- Otherwise, catching all cuts of size \( k \) requires at least \( n/k \) detection nodes.
- Sharp threshold impossible as well: catch all \( \varepsilon \)-cuts, but no cuts of size \( \varepsilon n \).
- Such a sharp cutoff requires at least \( n/2 \) detection nodes.

- Our Result: A detection set of size \( O(1/\varepsilon) \) that detects every \( \varepsilon \)-cut, and every reported cut has size at least \( \varepsilon n/2 \).
Geometry of Network Cuts

- Think of sensors as points in the plane.
- A linear cut is a line that partitions the point set.
- The point-line duality: point \((a,b)\) \(\Leftrightarrow\) line \((y = ax - b)\)

- It preserves above/below relationship:
  point \(p\) above line \(L\) \(\Leftrightarrow\) point \(L^*\) above line \(p^*\).
Geometry of Network Cuts

- A line $L$ is an $\varepsilon$-cut if the dual point $L^*$ lies above $\varepsilon n$ dual lines.
- Thus, the set of all linear $\varepsilon$-cuts is the region above the $\varepsilon n$ level (symmetrically, below the $(n - \varepsilon n)$ level).
- Imagine a polygonal curve (separator) made up of dual lines that lies between $\varepsilon n/2$ and $\varepsilon n$ levels.
- Then, the primal points corresponding to these lines form a detection set.
- Two issues:
  - Is there such a separator using just a few lines?
  - We don't know the cut line $L$. How will we decide that $L^*$ lies above the separator?
Complexity of a Separator

- A level can have $\Omega(n \log n)$ segments.
- Average complexity of a level is $\Theta(n)$.
- We want a separator of size roughly $1/\varepsilon$ (independent of $n!$).
- We construct a zig-zag path between levels $\varepsilon n/2$ and $\varepsilon n$.
  - Start at left, follow the edge until top level hit.
  - Reflect and follow until bottom level hit, reflect and continue.
- **Claim**: At least one zigzag separator path has $O(1/\varepsilon)$ segments. It lies between $\varepsilon n/2$ and $\varepsilon n$ levels.
Complexity of Separator

- Different zig zag paths are edge-disjoint.
- Total number of bends at most the number of vertices in the top and bottom levels.
- We choose two levels that each have $O(n)$ vertices, and are $\Omega(\varepsilon n)$ levels apart.
- By the averaging argument, at least one zigzag path will have $O(1/\varepsilon)$ segments.
- The dual points of these lines are our sentinels.
Detecting Cuts from Sentinels

- Base station computes the dual of the sentinel nodes only. (The arrangement formed by the separator lines.)
- Base station learns the dead/alive bit of each sentinel.
  - For each dead (live) sensor, we know that the dual of the cut $L^*$ must be above (below) the line.
  - In this example, $w_1, w_3, w_4$ are dead; others alive.
  - The intersection of these halfspaces gives a convex cell.
  - If this cell is above the separator, we declare an $\varepsilon_n$.
  - Otherwise, it's a false alarm.
Sample Sentinel Sets

Uniform
\[ N = 5000, \ \varepsilon = 0.01 \]
No. of sentinels = 14

Non-uniform
\[ N = 5000, \ \varepsilon = 0.01 \]
No. of sentinels = 14

US-census data
\[ N = 5000, \ \varepsilon = 0.01 \]
No. of Sentinels = 12
Scalability with Network Size

$\varepsilon = 0.01$

Size of Sentinel Set vs Network Size (X 1000)
Scalability with $\epsilon$

$N = 5000$

Size of Sentinel Set

uniform
non-uniform
US-Census

epsilon
Sentinels vs. Random Sampling

• Two natural random sampling schemes.
• Choose as many random nodes as our sentinel schemes.

• Random Sampling
  – $k$ nodes chosen uniformly at random.
  – Report if more than $\varepsilon k$ sentinels dead.

• Radial Sampling
  – $k$ directions chosen at random, and for each choose the $\varepsilon n$ extreme vertex.
  – Report if any of these $k$ dies.
False Positives

- Generated 250 cuts by picking points randomly between levels 1 and $\varepsilon n/2$
- These cuts are all below the appr threshold, and should not be reported.
- Random and radial sampling schemes misreport some of them as cuts.

No False Positives in Sentinel Set
False Negatives

- 250 cuts by picking points randomly between level $\varepsilon n$ and $2\varepsilon n$
- These are all above the approximation threshold, and so should be reported.
- Random and radial sampling schemes failed to report some of them as cuts.

No False Negatives in Sentinel Set
Extensions, Future Directions

- Robustness to random sensor failures
  - Make $c$ independent sentinel sets, take the majority vote to detect cuts
  - Fast randomized algorithm to find sentinel set

- Future work
  - More flexible cut definitions:
    - More complex shape.
    - Majority but not all sensors destroyed.
  - Distributed detection.