NETWORK PROBE SELECTION PROBLEM – CS

Kunal Gupta
Bhor Verma

About our Project

Effective use of the monitoring tools demands an understanding of monitoring requirements that system administrators most often lack. Instead of a well-defined process of defining monitoring strategy, system administrators adopt an ad-hoc, manual, and intuition-based approach. This leads to inconsistent and inadequate data collection and retention policies. For this project, we propose a solution to adapt the monitoring level of a system component using probes.

Processes involved

1. Setting Objectives & Mathematical Formulation
2. Converting to Matrix Form
3. Solving LP by implementation of solving techniques

1.

Mathematical Formulation

Understanding the problem, setting objectives
and coverting it to a mathematical one.

Problem Statement

The key idea is to send probes in the network, infer the system state from the probe results, and adjust the monitoring of system components based on the inferred system state. The key objective of the problem is:

» Probe selection - how to select the right set of probes such that criticality of components can be correctly estimated and appropriate recommendations can be provided for each component.

Assumptions

Network is like a undirected graph, with each node in graph denoting a node in the network.
Given,

- List of probes. (We can select out of these)
- List of nodes monitored by each node.
- Minimum number of probes by which each node has to be monitored. (For data consistency check)

Notations

Let the set of all probes be *P* and p represent any probe in *P*. *N* is the total number of probes.

Then,

- *nodes(p)* represents all the nodes that *p* can monitor,
- *C(p)* is the cost associated with probe p, where cost is defined as length of the probe = |*nodes(p)*|,
- and *c* is the minimum number of probes that monitor each node.

Conditions

Condition 1
The average number of probes probing each node:
$$\frac{\left(\sum_{p\in P_{sel}}|nodes(p)|\right)}{N}$$
should be maximized.

Condition 2
The average length of a probe:
$$\frac{\left(\sum_{p\in P_{sel}}|nodes(p)|\right)}{|P_{sel}|}$$
should be minimized.

Condition 3
Every node must have at least one probe passing through it.
$$\bigcup_{p\in P_{sel}}\left(\left(nodes(p)\right)\right) = N$$

Condition 4
In order to minimize the probe traffic, the number of selected probes:
$$|P_{sel}|$$
should be minimized.

Objective Function

We use LP formulation of the standard set cover with a slight modification to incorporate the desired constraints. The modification is required as the set cover problem requires each element to be covered in the selected subsets only once. Whereas, our requirement is that the node should be probed multiple times (*condition 1*). That is, each element in the universe must get covered at least c times in selected subsets.

Objective Function

We propose the following linear program that includes the four stated objectives in the form of constraints and objective functions:

$$min \sum_{p \in P} x_p \times C(p)$$

$$s.t. \sum_{p \in P} \left(x_p \times C(p) \right) \geq c \quad \forall n \in N$$

$$x_p \in \{0,1\} \quad \forall p \in P$$

$$\min \sum_{p \in P} x_p \times C(p)$$

$$s.t. \sum_{p \in P} \left(x_p \times C(p) \right) \geq c \quad \forall n \in N$$

$$x_p \in \{0,1\} \quad \forall p \in P$$

In the formulation,

- $x_p = \begin{cases} 1, & probe\ p\ is\ selected \\ 0, & otherwise \end{cases}$
- $p_n = \begin{cases} 1, & n \in nodes(p) \\ 0, & otherwise \end{cases}$

$$\min \sum_{p \in P} x_p \times C(p)$$

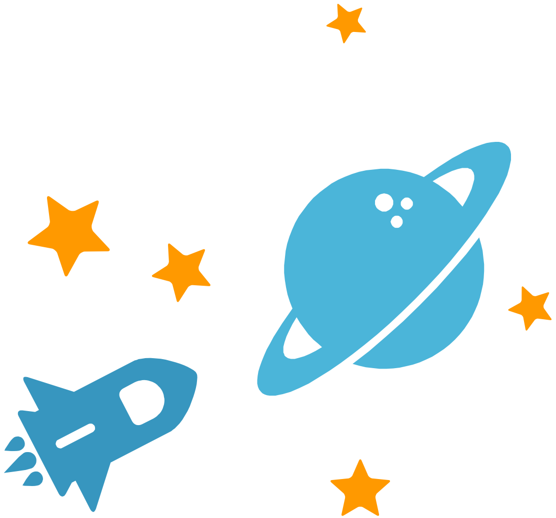
$$s.t. \sum_{p \in P} \left(x_p \times C(p) \right) \geq c \quad \forall n \in N$$

$$x_p \in \{0,1\} \quad \forall p \in P$$

The objective function tries to minimize the cost of all selected probes. It satisfies the 3rd and the 4th requirements. The constraint enforces that each node n in probe *p* should get probed at least *c* times, thereby satisfying the second requirement (average coverage). *c* is tuned based on the system requirements.

2.

Coversion to Matrix Form



Observation

p_n can be represented by a matrix $\mathrm{F}_{p\times n}$ having binary values. Thus,

$$p_n = F[p][n]$$

Hence, $nodes(p)$ is the pth row of F and so,

$$C(p) = \sum_{n\in N} F[p][n]$$

Now, $C(p)$ can be aggregated into a single vector of P dimensions just like x_p:

$$C = [C(p)]_{p\times 1}$$

$$x = [x_p]_{p\times 1}$$

Reformulation

From the previous observations we can reformulate our problem in the following form:

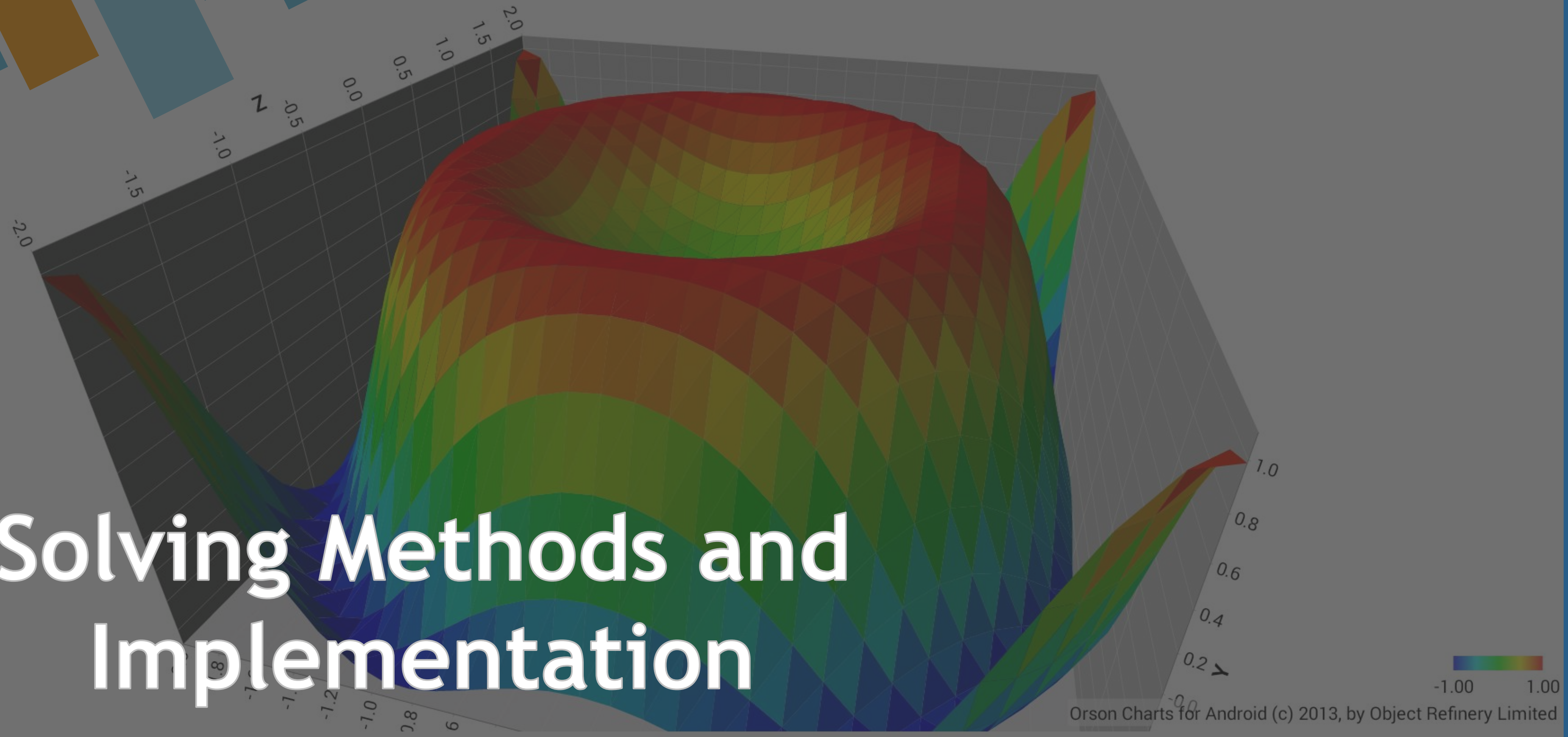
$$\min C^T x$$

$$s.t., F^T x \geq k$$

$$where\ k\ is\ c \cdot \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n \times 1}$$

$$\&\ each\ dimension\ of\ x\ can\ take\ values \in \{0, 1\}$$

Solving Methods and Implementation



Solving Methods

Base on the type of final LP we got after final reformulation, we found the following methods to solve it:

- BALAS
- Interior Point
- Lagrange and Sub-gradient

BALAS Algorithm

BALAS uses the Branch and Bound technique to eliminate many cases by eliminating whole subtrees of possible solution tree.

Click here to view BALAS.

THANKS!