Optimal Questionnaires for Screening of Strategic Agents
Anuj S. Vora, Ankur A. Kulkarni (Systems and Control Engineering, IIT Bombay)
anujvora@iitb.ac.in, kulkarni.ankur@iitb.ac.in

Background
- Health inspector encounters travellers
- Limited resources, not all travellers tested
- Identify susceptible travellers from their travel history
- However, people have a tendency to misreport their true travel history, due to stigma, inconvenience due to testing and quarantine protocols

Definitions
- **Sender graph** - Graph $G^n = (X^n, E)$, where $(x, y) \in E$ if $w_n(x, x, \lambda) \leq w_n(y, y, \lambda)$ or $w_n(y, y, \lambda) \leq w_n(x, y, \lambda)$.
- **λ-partition of a set** $\mathcal{P}^n$ is the largest subset of $\mathcal{P}^n$ which is an independent set in $G^n$.
- **Rate of information extraction** - For a strategy $g_n$, the rate of information extraction is
  \[ R(g_n) = \left( \sum_{\lambda \in \Lambda} P(\lambda) \min_{s \in \mathcal{P}(g_n, \lambda)} |D(g_n, s)| \right)^{1/n}. \]

Formulation of problem as a Stackelberg game
- We consider a Stackelberg game with the receiver as the leader.
- In a Stackelberg equilibrium, the strategies are
  \[ g^*_n = \arg \max_{g_n} \sum_{\lambda \in \Lambda} P(\lambda) \min_{s \in \mathcal{P}(g_n, \lambda)} |D(g_n, s)|. \]
- The best response set of the sender $\lambda$ is $\tilde{\sigma}(g_n, \lambda)$, where
  \[ \tilde{\sigma}(g_n, \lambda) = \left\{ s^*_n : X^n \rightarrow X^n \mid s^*_n(g_n \circ s^*_n(x), x, \lambda) \right\}. \]
- Assuming a pessimistic receiver, we incorporate the minimization over $\tilde{\sigma}(g_n, \lambda)$.

Example
- Let $\mathcal{X} = \{0, 1, 2\}$ and $\mathcal{A} = \{h, d\}$, where $h$: honest, $d$: dishonest.
- Let $P_A(h) = 1/3$ and $P_A(d) = 2/3$.
- Let the utility of the sender $h$ be
  \[ u^h(x, x, h) > u^h(x', x, h) \quad \forall x' \in \mathcal{X}, x' \neq x. \]
- For the sender type $\lambda = d$, the utility is
  \[ u^d(0, 0, d) = 0, u^d(1, 0, d) = 1, u^d(2, 0, d) = -1, u^d(0, 1, d) = 1, u^d(1, 1, d) = 0, u^d(2, 1, d) = -1, u^d(0, 2, d) = 1, u^d(1, 2, d) = 1, u^d(2, 2, d) = 0. \]
- Let $\mathcal{N} = 1$. Suppose $g(i) = i \forall i \in \mathcal{X}$, i.e., $C = \{0, 1, 2\}$.
- The best response of the sender $h$ is $s^h(x) = x$ for all $x \in \mathcal{X}$.
- For the sender type $d$, $\tilde{\sigma}(g, d) = \{ \tilde{s}^d(x) \}$, where
  \[ \tilde{s}^d(i) = \begin{cases} 1 & i = 0 \\ 0 & i = 1, 2 \end{cases}, \quad \tilde{s}^d(i) = \begin{cases} 1 & i = 0 \\ 0 & i = 1, 2 \end{cases}. \]

Main insight
- To get correct information, ask neither too many questions nor too less questions.

Conclusion
- We posed and studied a problem of information extraction from strategic receivers.
- Interesting observations even for a simple type of questionnaire.
- We characterized the optimal questionnaires for the receiver.
- We also derived bounds on the rate of information extraction, for finite $n$ and for $n$ growing to infinity.