Conditional Value-at-Risk Based Approaches to Robust Network Flow, Connectivity and Design Problems

Baski Balasundaram, Oklahoma State University
Vladimir Boginski, University of Florida
Stan Uryasev, University of Florida
Sergiy Butenko, Texas A&M University

U.S. DEPARTMENT OF ENERGY
Office of Science

Project Summary

- **Objective:** Develop models and algorithms for network flow, network design, and connectivity problems under uncertainty with the aim of obtaining robust solutions to the problems
- **Uncertainty:** Probabilistic node and arc failures modeled using uniform random graphs, random graphs of given expected degree sequence, and other models based on preferential attachment
- **Robustness:** By bounding or minimizing the conditional value-at-risk (CVaR) of an appropriately designed loss function, which quantifies losses as a function of decisions made under uncertainty
- **Problems investigated:** Minimum cost flows (including its special cases such as shortest paths, maximum flows, circulation), minimum spanning *k*-cores, minimum spanning *r*-robust *k*-clubs, and critical node detection
- **Research tasks:** Model development; theoretical study and algorithm design; large-scale implementation, testing and validation

Uncertainty in the network structure

- Underlying network is assumed to be uncertain; We assume that an edge exists with some probability quantified by random graph models of *given expected degree sequence*
- Given a sequence $w = (w_1,\ldots,w_n)$, a random graph $G(w)$ is defined where the probability of an edge between $i, j \in V$ is given by:

$$p_{ij} = \frac{w_i w_j}{\sum_{k \in V} w_k}$$

- Erdos-Renyi random graph model $G(n, p)$ is obtained from G(w) by choosing $w_i = np$
- $G(w)$ represents a power-law random graph model when sequence w obeys a power-law
- Power-law model is particularly interesting as network models of many natural and engineered complex systems have been empirically shown to exhibit power-law degree distributions
- Models of preferential attachment, and duplication models describing power-law graphs will also be investigated

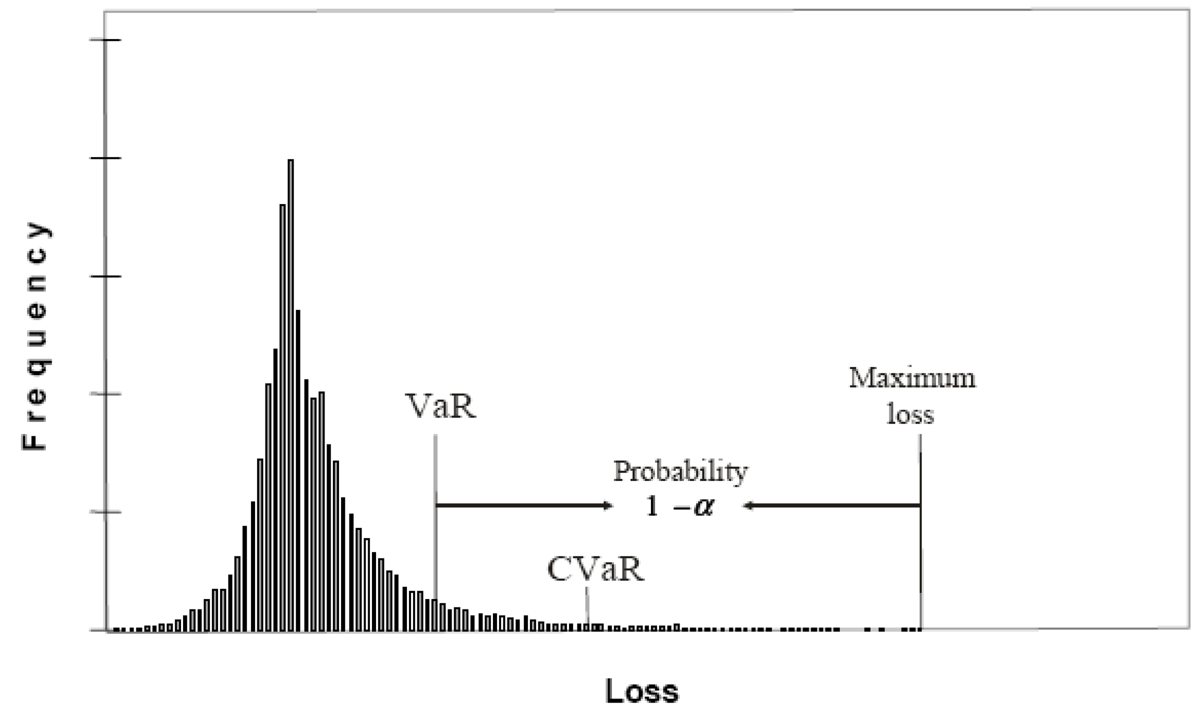
Robustness Via Conditional Value-at-Risk

- *Loss function:* Suppose $L(\boldsymbol{x},\boldsymbol{Y})$ quantifies losses associated with decision vector $\boldsymbol{x}$ and random vector $\boldsymbol{Y}$ of uncertain parameters
- *α-Value-at-Risk* is the α-percentile of the distribution associated with the loss function which depends on $\boldsymbol{x}$

$$\alpha\text{-VaR}[L(x,Y)] = \min\{q : P\{L(x,Y) \le q\} \ge \alpha\}$$

- *α-Conditional Value-at-Risk* is the expectation of losses exceeding *α-VaR,* the mean of the α tail distribution of $L(\boldsymbol{x},\boldsymbol{Y})$,

$$\alpha\text{-CVaR}[L(x,Y)] = E[L(x,Y) | L(x,Y) \ge \alpha\text{-VaR}[L(x,Y)]]$$



- A decision $\boldsymbol{x}$ that minimizes (or limits) *α-CVaR* of $L(\mathbf{x},\mathbf{Y})$ minimizes (or limits) the average losses in the worst 1-α percentage of cases, and is considered a robust solution
- CVaR aggregates various losses under uncertainty into a single coherent measure of downside risk and is more conservative than VaR
- CVaR is convex in $\boldsymbol{x}$ if $L(\boldsymbol{x},\boldsymbol{Y})$ is convex in $\boldsymbol{x}$, which is desirable if CVaR is minimized or bounded from above, inside an optimization framework
- Explicit distribution for $L(\boldsymbol{x},\boldsymbol{Y})$, explicit expression for CVaR, or computation of VaR are not needed, instead the following convex function can be used in place of CVaR:

$$F_\alpha(x,\zeta) = \zeta + \frac{1}{1-\alpha} E\left[(L(x,Y) - \zeta)^+\right]$$

- Since CVaR is an expectation, higher order moments can be used to quantify risk of losses under uncertainty for heavy tailed distributions

The Minimum Spanning k-Core Problem

- The proposed idea to employ CVaR to obtain robust solutions to optimization problems can be illustrated using a novel network design problem studied in this project, based on the concept of *k*-cores introduced by Seidman in 1983 for social network analysis

> **Definition (Seidman, 1983)**
> *Given a simple undirected graph $G = (V, E)$, for $S \subseteq V$, the induced subgraph $G[S]$ is called a k-core if the minimum degree in $G[S]$ is at least k.*

- One can add edges to the network (design it) so that it results in a *k*-core; Appropriate choice of parameter *k* can guarantee desired diameter and vertex connectivity; This leads to *the minimum spanning k-core problem*

> **Definition**
> *Given $V = \{1,\ldots,n\}$, c_{ij} for each distinct i, j pair and a fixed positive integer k, identify a minimum cost set of edges E^* to be created so that the graph $G = (V, E^*)$ is a k-core.*

- The minimum spanning *k*-core problem is polynomial-time solvable using generalized graph matching techniques
- Suppose edges have probabilities of survival/failure, hence a spanning *k*-core could cease being one if some edges we chose, failed; Consider the loss function,

$$L(x,Y) = \sum_{v \in V} \Big(k - \sum_{e \in \delta(v)} x_e Y_e\Big)^+$$

which measures the cumulative degree deficiency under uncertainty where x_e is a binary variable indicating if e is chosen to be included, and Y_e is an indicator random variable for edge e surviving

- The CVaR constrained minimum spanning *k*-core problem under uncertainty can be formulated as:

> **Formulation (CVaR constrained)**

$$\begin{aligned} \min \sum_{e \in E} c_e x_e & & \\ \sum_{e \in \delta(v)} x_e &\ge k, && v \in V \\ \zeta + \frac{1}{1-\alpha} \sum_{s=1}^{|\Omega|} p_s\left[(L(x,y^s) - \zeta)^+\right] &\le C && \\ \zeta &\in \mathbb{R} && \\ x_e &\in \{0,1\}, && e \in E \end{aligned}$$

- We expect to be able to develop a polynomial time separation algorithms since the original problem is polynomial time solvable; This is under investigation
- CVaR formulations of min cost flow problem and algorithms have been developed; In this case, the number of samples needed to estimate CVaR to desired accuracy is polynomial in the number of arcs

UF UNIVERSITY of FLORIDA | OKLAHOMA STATE UNIVERSITY | TEXAS A&M UNIVERSITY