ARPA-E GO COMPETITION CHALLENGE 1: SCORING

UPDATED: 2018-08-28

1. TERMINOLOGY

The following terminology will be utilized throughout the GO Competition and in this scoring document:

- ***Power system network model***: each hypothetical grid with defined topological structure and characteristics including, but not limited to, locations of generators, loads, transmission lines, transformers, equipment detail, control equipment, and limits.
- ***Scenario***: an operating instance in time on a power system network model. The scenarios define an instantaneous demand at each bus, renewable resource availability, and other temporary system conditions.
- ***Dataset***: A collection of power system network models and scenario data on those models. Challenge 1 will have four distinct datasets: C10D, C1TD1, C1TD2, C1FD (see Section 2 for more details).
- ***A scenario score*** (relevant to Division 1 and Division 2, see below) is calculated for each scenario of a power system network model.
- ***A power system network model score*** (relevant to Divisions 1 and Division 2) is calculated by taking the geometric mean across all scenarios associated with a network model.
- ***A dataset score*** (relevant to Division 1 and Division 2) is computed by taking the geometric mean of all power system network models in a given dataset.

2. GO COMPETITION DIVISIONS

The GO Competition will host four different "divisions" with separate leaderboards and scoring algorithms. Two will be focused on real-time optimization (with a 10 minute time limit per scenario) and two focused on offline optimization (with a 45 minute time limit per scenario). For each time period, one of the scoring divisions will focus on value of the objective function while the other will focus on algorithm robustness measured via performance profiles.[1,2] The Eligible Entrants in the Top 10 of each division will receive $100,000 (up to $400,000) in either prize money or grants for follow-on research approved in writing by ARPA-E (for more details

---

[1] Elizabeth D. Dolan and Jorge J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, 91(2):201-213, 2002. DOI: 10.1007/s101070100263.

[2] Nicholas Gould and Jennifer Scott, "A Note on Performance Profiles for Benchmarking Software," *ACM Transactions on Mathematical Software* (TOMS) 43.2 (2016): 15.

on Open Entrants versus Proposal Entrants, see Section 4). Prizes (grants for Proposal Entrant awardees), for placing in multiple divisions, are additive. Figure 1 depicts algorithm scoring by division. Scoring for all four divisions is discussed below.

Figure 1. Breakdown of the Scoring Divisions and Prize/Grant Money for Eligible Entrants.

A. Divisions 1 and 2: Objective Function Scoring:

A particular Entrant's solution score will be referred to as score c. A solution score exists for each network model and scenario; however, without loss of generality, we will suppress indices for model and scenario in the following discussion.[3] The scenario score is defined as:

$$c = \sum_{g \in G} C_g(p_g) + \delta c^{\sigma} + \frac{(1-\delta)}{|K|} \sum_{k \in K} c_k^{\sigma} \qquad (1)$$

where G is the set of generators and K is the set of contingencies. $C_g(p_g)$ is the cost to produce real power at generator g in the base case[4] (in each network model and under each scenario).[5] Note that $C_g(p_g)$ will be represented by a piecewise linear cost formulation with up to ten segments; the marginal costs are monotonically non-decreasing from segment to segment. c^{σ} represents the total penalty costs for violations in the base case (pre-contingency) and c_k^{σ} represents the total penalty costs for violations that occur in post-contingency k. In the official formulation, σ is used as a slack variable to relax a variety of constraints for both pre-contingency and post-contingency operations: real and reactive power node balance equations, line current ratings, and transformer power ratings. δ is the assigned probability that no contingency occurs and is used to weight violations in the base case and contingencies. See ARPA-E DE-FOA-0001952 Appendix A2 and the full formulation on the website:

[3] Given a network model, $m \in M$, a scenario, $s_m \in S_m$, and an Entrant, $\tau \in \mathrm{T}$, the resulting score for that particular instance is represented as $c_{m,s_m,\tau}$. Note that $s_m \in S_m$ is indexed by model $m \in M$ as different models may have different number of scenarios. Therefore, each model m can have its own corresponding set of scenarios, $s_m \in S_m$.

[4] The costs to redispatch the fleet of generators in response to a contingency is not included in the objective.

[5] See Appendix A2 of ARPA-E DE-FOA-0001952 and the full formulation on the website: https://gocompetition.energy.gov/challenges/challenge-1/formulation for a complete coverage of nomenclature.

https://gocompetition.energy.gov/challenges/challenge-1/formulation for the SCOPF mathematical formulation.

Thus, the score is the objective function value under a given network model, m, and scenario, s_m, which is defined to be the addition of the total cost and linearly penalized slack variables, corresponding to real and reactive power nodal imbalances as well as branch overloading. The costs of these violations are captured by the penalty prices (see λ in the official formulation online) and will be provided as part of each power system network model and based on the prices used in real industry markets.[6] As discussed, Division 1 will evaluate this score after t_1 =10 minutes on the competition platform, while Division 2 will do so with set-points returned after t_2 =45 minutes of evaluation. Algorithms will be assigned a penalty score (for infeasibility) of c_{m,s_m}^{inf} for a given power system network model m and a given scenario s_m if they do not return a solution in the given amount of time or the solution is infeasible to the official SCOPF formulation. Note that the official SCOPF formulation includes a relaxation of the nodal real/reactive power balance constraints and branch limits with the inclusion of slack variables and penalty prices on those slacks; feasible solutions for this relaxed SCOPF problem are scored based on (1). Infeasible solutions are those that violate constraints that are not relaxed in the SCOPF formulation (for example, power flow equations or a voltage limit). The score given to any solution that is infeasible, c_{m,s_m}^{inf}, to the relaxed SCOPF problem is the maximum of two other scores: c_{m,s_m}^{max} and c_{m,s_m}^{slack}. c_{m,s_m}^{max} is the largest cost for a feasible solution achieved by any Entrant, $\tau \in \mathrm{T}$: $c_{m,s_m,\tau}$. c_{m,s_m}^{slack} is the cost when satisfying all ***net load*** by the slack variables directly. Equations (2)-(4) represent these relationships. **Note that the cost of c_{m,s_m}^{inf} will only be applied during the Trial 1 Event, Trial 2 Event, and the Final Event (the competition) for infeasible solutions; the continuously updated leaderboards (outside of the Trial Events and the Final Event) will reflect a cost of c_{m,s_m}^{slack} for infeasible solutions.** Note that (3) is a simplified expression of the actual c_{m,s_m}^{slack}. In (3), n represents a segment in a piecewise linear penalty cost function for violations (constraint relaxations), λ_n^P and λ_n^Q represent the corresponding penalty prices for that segment, and $D_{i,m,s_m,n}^P$ and $D_{i,m,s_m,n}^Q$ represent the amount of demand at a bus that is attributed to that segment's length. Together, (3) provides a feasible solution when serving all of the load via the slack variables alone.

$$c_{m,s_m}^{max} = \max_{\forall\tau}\{c_{m,s_m,\tau}\} \tag{2}$$

$$c_{m,s_m}^{slack} = \textstyle\sum_{\forall i\in I, n\in N}\left(\lambda_n^P D_{i,m,s_m,n}^P\right) + \sum_{\forall i\in I, n\in N}\left(\lambda_n^Q D_{i,m,s_m,n}^Q\right) \tag{3}$$

$$c_{m,s_m}^{inf} = \max\{c_{m,s_m}^{max}, c_{m,s_m}^{slack}\} \tag{4}$$

On page 4, Figure 2 shows a visual representation of two notional algorithms, with Algorithm 1 winning Division 1, and Algorithm 2 winning Division 2. Entrants will be able to submit their executable program (source code) with the ability to adjust their algorithmic approach based on division; there will be an input parameter that will reflect the selected division.

[6] California ISO. Market Parameter Settings for MRTU Market Launch. February 2009.

Once an Entrant's executable program has a score for a given network model, $m \in M$, for each scenario, $s_m \in S_m$, the power system network model score is computed by taking a geometric mean over the number of scenarios for that network model:

$$Score_{m,\tau} = \sqrt[|S_m|]{\prod_{s_m} c_{m,s_m,\tau}} \quad (5)$$

Similarly, the total dataset score, for Entrant $\tau \in \mathrm{T}$, is a geometric mean over the power system network model scores:

$$Score_{\tau} = \sqrt[|M|]{\prod_{m} Score_{m,\tau}} \quad (6)$$

Winners of Division 1 and Division 2 will be determined by their rankings on the dataset score for P1FD, evaluated during the GO Competition Final Event.

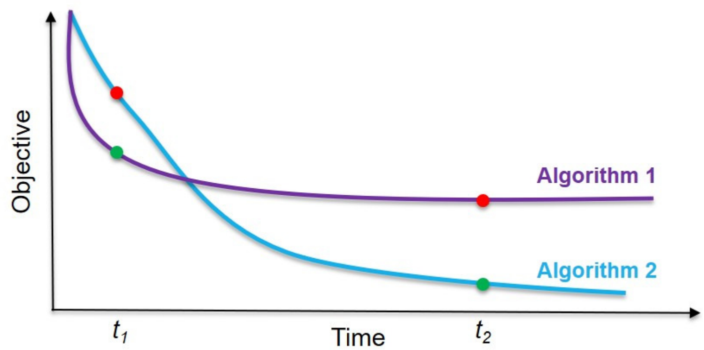

Figure 2. Algorithm Performance over Time.

B. Divisions 3 and 4: Performance Profile Scoring:

The creation of performance profiles[7] is a method to compare algorithmic approaches by constructing the cumulative distribution function of a particular performance metric.[8] First, any Entrant's approach for model m and scenario s_m, which did not produce a feasible solution to the relaxed SCOPF problem specified in Appendix A2 of ARPA-E DE-FOA-0001952 and the full formulation on the website: https://gocompetition.energy.gov/challenges/challenge-1/formulation, will not be considered within the performance profile ranking. Next, we define $r_{m,s_m,\tau}$ as the ratio of Entrant τ's score relative to the best score for model m and scenario s_m. If the Entrant was unable to produce a feasible solution for model m and scenario s_m, then no value of $r_{m,s_m,\tau}$ is given, which reflects that such an infeasible answer is not counted. $r_{m,s_m,\tau}$ is then defined as:

$$r_{m,s_m,\tau} = \frac{c_{m,s_m,\tau}}{\hat{c}_{m,s_m}} \quad (7)$$

[7] Elizabeth D. Dolan and Jorge J. Moré "Benchmarking optimization software with performance profiles," *Mathematical Programming*, 91(2):201-213, 2002. DOI: 10.1007/s101070100263.

[8] In the literature, this performance metric is typically a ratio of solve-times; however, here we construct it as a ratio of total dispatch cost.

where: $\hat{c}_{m,s_m} = \min_{\tau}(c_{m,s_m,\tau})$, which represents the lowest objective function found on model m for scenario s_m across all submitted algorithms from Entrant τ. The best performing algorithm on model m for scenario s_m has $r_{m,s_m,\tau} = 1$ while all others have $r_{m,s_m,\tau} > 1$. Next, we define $k(f, r_{m,s_m,\tau})$, a counter variable that gives 1 if $r_{m,s_m,\tau}$ is less than or equal to f ($f \geq 1$) and 0 otherwise. Finally, we can define:

$$p_\tau(f) = \frac{\sum_{\forall m,s_m} k(f, r_{m,s_m,\tau})}{\sum_{\forall m} |s_m|} \quad (8)$$

For a given value of f and for Entrant τ, $p_\tau(f)$ reflects the fraction of scenarios (for all models) in which the Entrant's algorithmic approach performed within factor f of the best solution submitted across all Entrants. For example, $p_\tau(1)$ gives the fraction of scenarios where Entrant τ produced the lowest cost, while $p_\tau(\infty)$ gives the fraction of scenarios (for all models) where Entrant τ produced a feasible solution to the relaxed SCOPF problem (see Appendix A2 of ARPA-E DE-FOA-0001952 and the full formulation on the website: https://gocompetition.energy.gov/challenges/challenge-1/formulation).

A "performance profile" is then the function above considered over a range of f (frequently plotted on a semi-(binary) logarithmic scale). Figure 3 shows example performance profiles for 12 different algorithms.

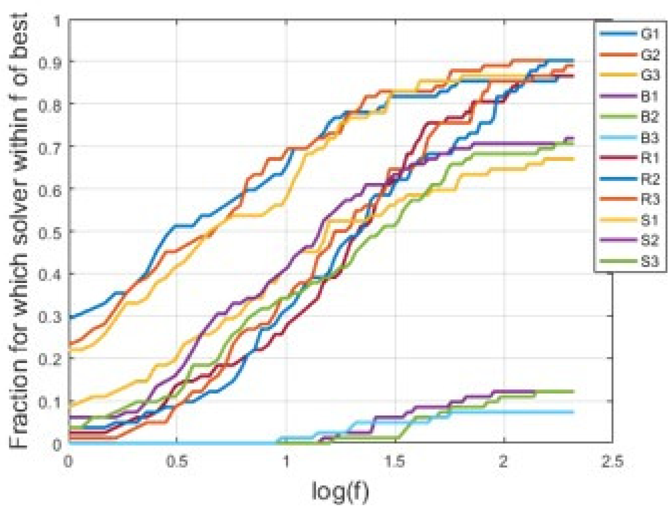

Figure 3. Performance Profiles across 12 Algorithms.

To rank each Entrant's submission, we rank each submission according to the area under each performance profile curve, integrated up to a chosen f^{max}. However, it has been pointed out that ranking based on the integral of each performance profile over a certain range unfairly penalizes algorithms that are in "second place" across many scenarios. [9] Therefore, we amend the ranking as follows: all algorithms' performance profiles are drawn and the one with the largest integral is the "winner." That algorithm (the winner) is then removed from consideration

[9] Nicholas Gould and Jennifer Scott, "A Note on Performance Profiles for Benchmarking Software," *ACM Transactions on Mathematical Software* (TOMS) 43.2 (2016): 15.

and the remaining performance profiles are recalculated and redrawn to generate second place, and so on, successively.

For Division 3, any Entrant submission failing to produce a feasible solution (a solution to the relaxed SCOPF problem specified in Appendix A2 of ARPA-E DE-FOA-0001952 and the full formulation on the website: https://gocompetition.energy.gov/challenges/challenge-1/formulation) within the time limit for Division 3, i.e., t_1 =10, will not receive a score for $r_{m,s_m,\tau}$. Division 4 is constructed similarly, but with t_2 =45 minutes. The value of f^{max} will be announced on the GO Competition website around the time that the GO Competition Challenge 1 is launched.

Note that performance profile scores (Divisions 3 and 4) will only be assessed during the Trial 1 Event, Trial 2 Event, and the Final Event; the continuously updated leaderboards (outside of the Trial Events and Final Event) will only reflect lowest cost scores (Divisions 1 and 2).