Supplementary file: Anomaly Detection and Attribution in Networks with Temporally Correlated Traffic

Ido Nevat, Dinil Mon Divakaran, Sai Ganesh Nagarajan, Pengfei Zhang, Le Su, Li Ling Ko, Vrizlynn L. L. Thing

I. ANOMALY DETECTION USING THE CE METHOD

The generic CE algorithm is presented in Algorithm 1. Typically, for combinatorial optimization problems, the stopping criterion is the point when the difference in the objective value is less than a predefined minimum.

Algorithm 1 Cross Entropy Generic Algorithm

while stopping criterion ≠ TRUE do
1. Simulate candidate solutions from a parametrized sampling distribution.
2. Calculate the performance function score for each candidate.
3. A fixed number of best performing candidates are retained and they are denoted as the “elite samples”.
4. Use the “elite samples” to update the parameters of the sampling distribution in the next iteration.
end while

APPENDIX A

PROOF OF LEMMA 1

\[
\Lambda \left( S_{1:T}^{(1,K)} \right) = \frac{\Pr \left( S_{1:T}^{(1,K)} \right) \Pr \left( S_{1:T}^{(K)} | H_0 \right)}{\sup \Theta \Pr \left( S_{1:T}^{(1,K)} \right) \Pr \left( S_{1:T}^{(K)} | H_1 \right)} = \frac{\prod_{k=1}^{K} \Pr \left( S_{1:T}^{(k)} | H_0 \right)}{\sup \Theta \prod_{k=1}^{K} \Pr \left( S_{1:T}^{(k)} | H_1 \right)} = \sup \Theta \left\{ \prod_{k \in K_u} \left\{ \prod_{i=1}^{T} \Pr \left( S_{i:T}^{(k)} | Q^i_0 \right) \prod_{i=1}^{T} \Pr \left( S_{i:T}^{(k)} | S_{i-1:T}^{(k)} \right) \right\} \right\} \times \sup \Theta \left\{ \prod_{k \in K_a} \left\{ \prod_{i=1}^{T} \Pr \left( S_{i:T}^{(k)} | Q^i_0 \right) \prod_{i=1}^{T} \Pr \left( S_{i:T}^{(k)} | S_{i-1:T}^{(k)} \right) \right\} \right\}.
\]

APPENDIX B

PROOF OF LEMMA 2

To obtain the test statistic for the LRT, we need to calculate the following joint density, where we generically state the transition matrix as \( P \).

\[
\begin{align*}
\Pr (Z_{1:T}, F_{2:T} | P) &= \Pr (Z_1, Z_2, \ldots, Z_T, F_2, F_3, \ldots, F_T | P) \\
&= \Pr (Z_2, Z_T, F_2, \ldots, F_T | P) \Pr (Z_1, P) \\
&= \prod_{t=2}^{T} \Pr (Z_t | Z_{t-1}, F_t) \Pr (F_t | Z_{t-1}, P) \Pr (Z_1) \\
&= \prod_{t=2}^{T} \Pr (Z_t | Z_{t-1}, F_t) \Pr (F_t | Z_{t-1}, P) \Pr (Z_1) \\
&= \prod_{t=2}^{T} \left( \prod_{i=1}^{J} Z_t^{(i)} \prod_{k=1}^{J} \left[ \frac{P_{i,k}^{(F_t)}}{F_{i,k}^{(F_t)}} \right] \right) \times \Pr (Z_1).
\end{align*}
\]

Next, utilizing this results we obtain a compact expression for the test statistic:

\[
\Lambda (S_{1:T}) = \log \frac{\Pr \left( S_{1:T}^{(1,K)} = s_{1:T}^{(1,K)} \right) | H_0}{\sup_{Q_a} \Pr \left( S_{1:T}^{(1,K)} = s_{1:T}^{(1,K)} \right) | H_1} = \log \frac{\Pr (Z_{1:T} = z_{1:T}, F_{2:T} = f_{2:T} | Q_a)}{\Pr (Z_{1:T} = z_{1:T}, F_{2:T} = f_{2:T} | Q_a)}
\]

\[
= \log \frac{\Pr (Z_t | Q_a) \prod_{t=2}^{T} \left( \prod_{i=1}^{J} Z_t^{(i)} \prod_{k=1}^{J} \left[ \frac{Q_{i,k}^{(F_t)}}{F_{i,k}^{(F_t)}} \right] \right)}{\prod_{t=2}^{T} \left( \prod_{i=1}^{J} Z_t^{(i)} \prod_{k=1}^{J} \left[ \frac{Q_{i,k}^{(F_t)}}{F_{i,k}^{(F_t)}} \right] \right)}
\]

\[
= \log \Pr (Z_t | Q_a) - \log \Pr (Z_t | Q_a) + \sum_{t=2}^{T} \log \left( \prod_{i=1}^{J} Z_t^{(i)} \prod_{k=1}^{J} \left[ \frac{Q_{i,k}^{(F_t)}}{F_{i,k}^{(F_t)}} \right] \right)
\]

\[
= \log \Pr (Z_t | Q_a) - \log \Pr (Z_t | Q_a) + \sum_{t=2}^{T} \sum_{k=1}^{J} \left[ \log |Q_a|_{i,k} - \log |F_t|_{i,k} \right]
\]

\[
= \log \Pr (Z_t = z_t | Q_a) - \log \Pr (Z_t = z_t | Q_a) + \sum_{t=2}^{T} \sum_{k=1}^{J} \left[ \log |Q_a|_{i,k} - \log |Q_a|_{i,k} \right].
\]