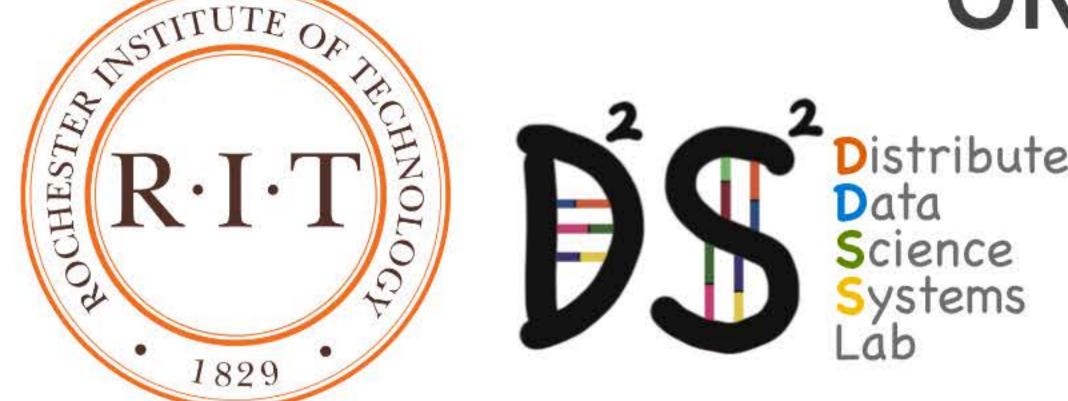# ONE-NAS: An Online NeuroEvolution based Neural Architecture Search for Time Series Forecasting

Zimeng Lyu, Travis Desell

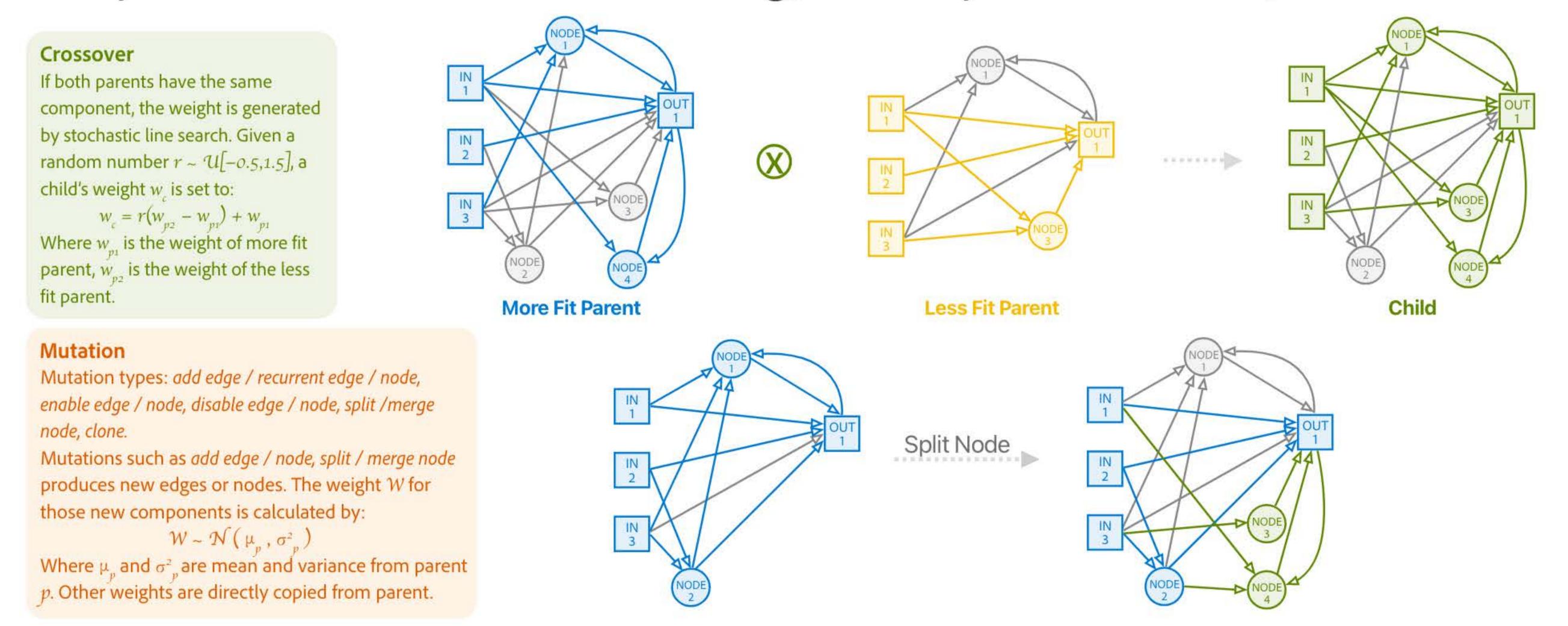Rochester Institute of Technology, Rochester, New York, USA

Motivation

- Time series forecasting (TSF) is important in many domains of data science
- Usually time series data is non-stationary
- TSF ML models typically train on historical data and apply on unseen data
- Models usually suffer from **data drift** or **concept drift** problem
- Periodically retrain and redesign models takes significant resources

About ONE-NAS

- **First** algorithm capable of evolving RNNs **online** for TSF
- RNNs are generated & trained online **without** pre-training, predictions are made online
- Outperform Naive, moving average, and exponential smoothing prediction
- Use island repopulation strategy to speed up the evolution
- Asynchronous distributed scheduling, naturally load balanced, real-time

Crossover

If both parents have the same component, the weight is generated by stochastic line search. Given a random number $r \sim \mathcal{U}[-0.5, 1.5]$, a child's weight w_i is set to:

$$w_c = r(w_{p1} - w_{p2}) + w_{p2}$$

Where w_{p1} is the weight of more fit parent, w_{p2} is the weight of the less fit parent.

Mutation

Mutation types: *add edge / recurrent edge / node, enable edge / node, disable edge / node, split /merge node, clone.*

Mutations such as *add edge / node, split / merge node* produces new edges or nodes. The weight W for those new components is calculated by:

$$W \sim \mathcal{N}(\mu_p, \sigma_p^2)$$

Where μ_p and σ_p^2 are mean and variance from parent p. Other weights are directly copied from parent.

More Fit Parent ⊗ Less Fit Parent → Child

Split Node

Online NeuroEvolution Neural Architecture Search (ONE-NAS)

Worker 1 / Worker 2 / Worker 3 — Train

Main

Untrained Offspring O_t — Trained Offspring O_t

Generate RNN — *Crossover* ⊗ *Mutate*

Fitness Function

$$\mathcal{E}_t + O_t \xrightarrow[\text{Evaluate+Select}]{\mathcal{B}_{validate}} \mathcal{E}_{t+1}$$

Elite Population $\mathcal{E}_t$ — Elite Population $\mathcal{E}_{t+1}$

Island i — Island j — Island 1 — Island 2 — Island 3

Island 1 $\mathcal{E}_t$
Island 2 $\mathcal{E}_t$
... → Select → Global Best Genome → Online Prediction

New generated Offspring | Trained Offspring | Elite RNN | Crossover operator

https://github.com/travisdesell/exact

ONE-NAS Prediction Example

Expected / Predicted. X-axis: Samples (0–7000). Y-axis: P-avg (Normalized) 0.0–1.0

One-NAS vs. One-NAS Repopulation

Naive — Moving Average — Exponential Smoothing. Y-axis: MSE. Categories: OneNas, OneNAS Repopulation

One-NAS repopulation

Naive. Y-axis: MSE. Categories: Island 20 Extinct Freq: 200, Island 20 Extinct Freq: 350, Island 20 Extinct Freq: 500, Island 30 Extinct Freq: 200, Island 30 Extinct Freq: 500, Island 40 Extinct Freq: 100, Island 40 Extinct Freq: 350

Full Paper Link

RIT | Rochester Institute of Technology