PAUL SCHERRER INSTITUT


S. Li, et al. :: LSM :: Paul Scherrer Institut

# HIPAInterlock Survival Modelling Overview

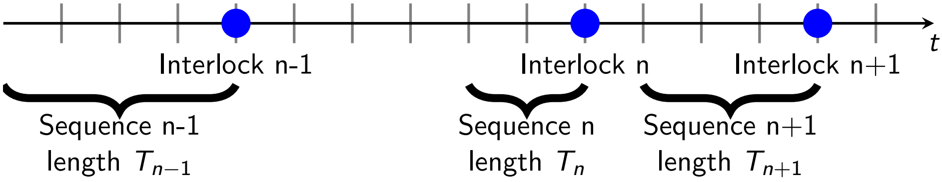
July 3, 2020

PAUL SCHERRER INSTITUT PSI

# Problem Formalisation

SDSC

- For each $t$, $\vec{x}_t$ is the time series of $N_{feature} = 359$ input channels
- For each $t$, $y_t$ is the observed time to next interlock (TTI)



# Seq2One model



- $\vec{x}^n_{t-\tau:t}$ is the window of input data of sequence $n$ from $t-\tau$ to $t$. It has shape $(\tau, N_{feature})$.
- $y^n_t$ is the observed TTI of sequence $n$ at time $t$ (i.e. time to the end of sequence)
- $Y$ is TTI random variable following Weibull distribution

$$Y \sim f_Y(y|\theta) = \frac{\beta}{\alpha}\left(\frac{y}{\alpha}\right)^{\beta-1} e^{-(y/\alpha)^{\beta}}, \quad y \geq 0, \quad \theta = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} \tag{1}$$

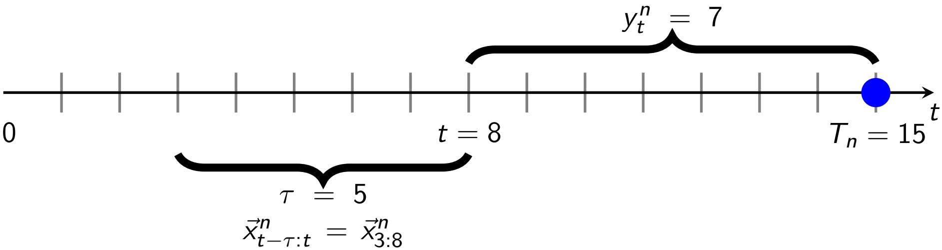
- $\theta = \begin{pmatrix}\alpha \\ \beta\end{pmatrix}$ is the output of our NN model $\boldsymbol{g}$

$$\theta = \boldsymbol{g}(\vec{x}|\vec{x}^n_{t-\tau:t}, y^n_t), \quad t = \tau, \ldots, T_n, \quad n = 1, \ldots, N_{seq} \tag{2}$$

- The model is trained by maximising the log likelihood

$$\boldsymbol{g} = \max_{\boldsymbol{g}} \mathcal{L} = \max_{\boldsymbol{g}} \sum_{n=1}^{N_{seq}} \sum_{t=\tau}^{T_n} \log Pr(Y = y^n_t|\theta^n_t) = \max_{\boldsymbol{g}} \sum_{n=1}^{N_{seq}} \sum_{t=\tau}^{T_n} \log f_Y(y^n_t|\boldsymbol{g}(\vec{x}^n_{t-\tau:t})) \tag{3}$$

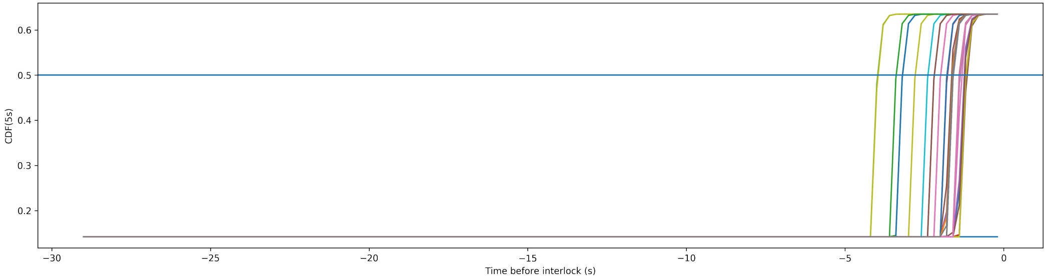
# Seq2One model



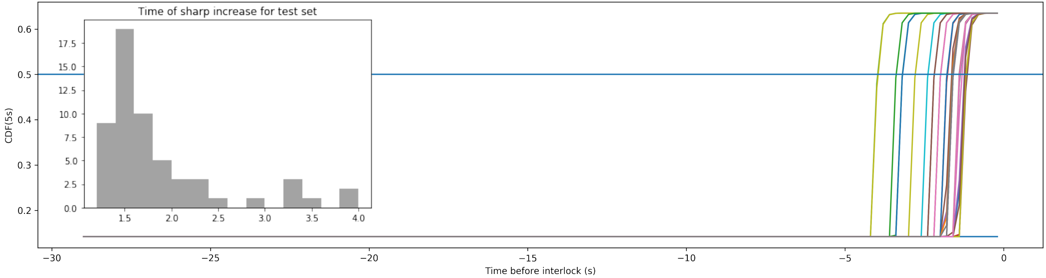
# Seq2One model



Result for $\tau = 5$ timesteps (1s), $T_n = 150$ timesteps (30s) for all $n$ in test set: 55/56 interlocks can be detected 1s $\sim$ 4s before interlock

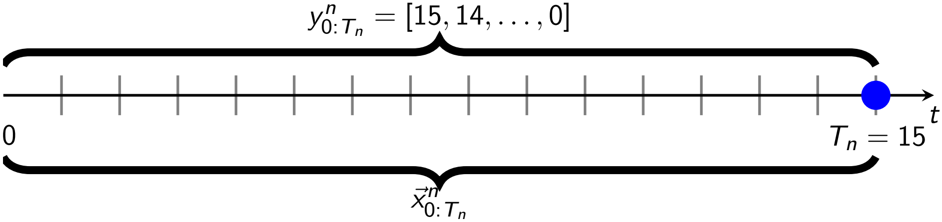


Result for $\tau = 5$ timesteps (1s), $T_n = 150$ timesteps (30s) for all $n$ in test set: 55/56 interlocks can be detected 1s $\sim$ 4s before interlock

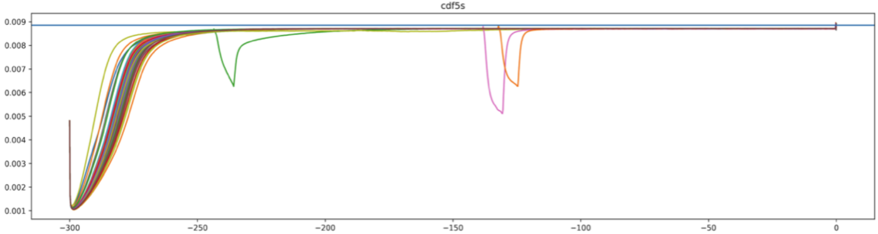


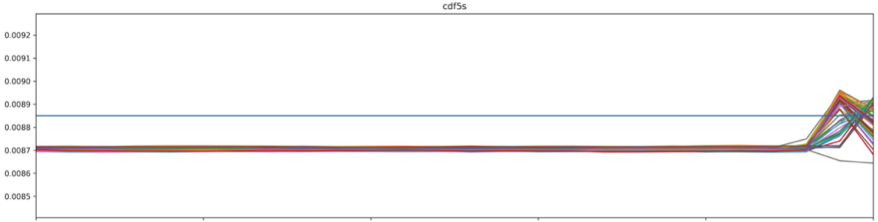
# Seq2Seq model

- $\vec{x}^n_{0:T_n}$ is the input data of sequence $n$. It has shape $(T_n, N_{feature})$.
- $y^n_{0:T_n}$ is the observed TTI array of sequence $n$. It has shape $(T_n, 1)$.
- $\theta^n_{0:T_n} = \begin{pmatrix} \alpha^n_{0:T_n} \\ \beta^n_{0:T_n} \end{pmatrix}$ is the output of our NN model
- In reality, train with various sub-sequences $\{\vec{x}^n_{t_1:t_2}, y^n_{t_1:t_2}\}$ with $t_1, t_2 \in [0, T_n]$



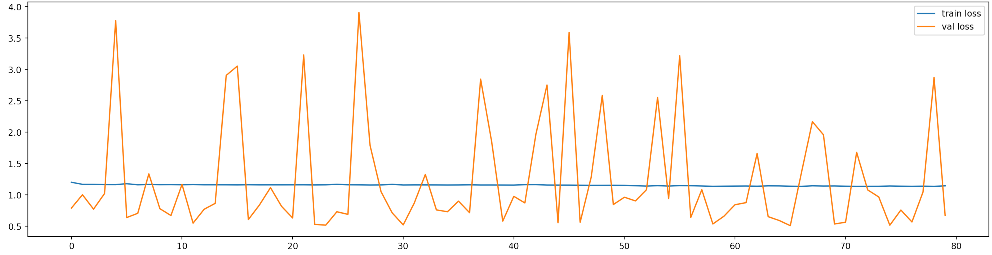
Result for $T_n = 1500$ timesteps (5min) for all $n$ in test set: 55/56 interlocks can be detected, but only 0.2s before and with tiny signals



# Outlook

- Seq2One: Longer time before interlock $\rightarrow$ imbalance problem
- Seq2Seq: Problem of loss function




- Involve $t$ (time that already runs without interlock) in the model

**Thanks to**

Andreas Adelmann
Jaime Coello
Perez Cruz Fernando
Anastasia Pentina
Davide Reggiani
Jochem Snuverink
Mélissa Zacharias