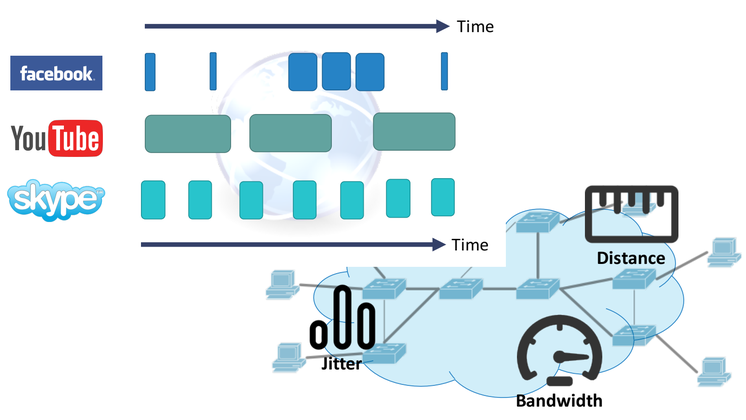
Applying Deep Learning to Network Traffic Identification and Categorization

Laboratory for Software Design & Analysis
established in 2005

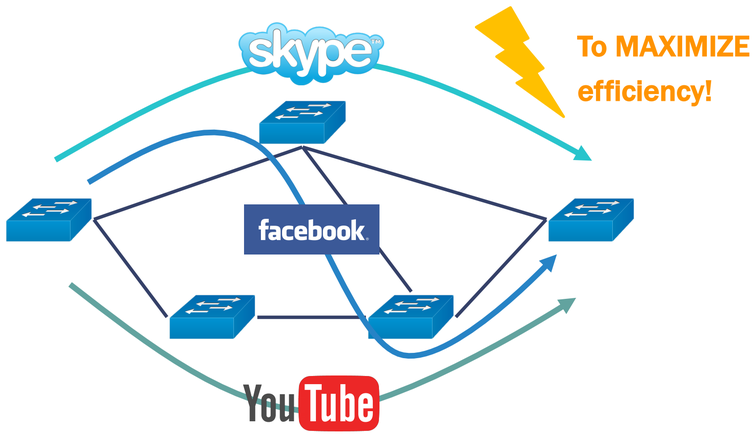
Pongsakorn U-chupala, Yasuhiro Watashiba,
Kohei Ichikawa and Iida Hajimu

Motivation

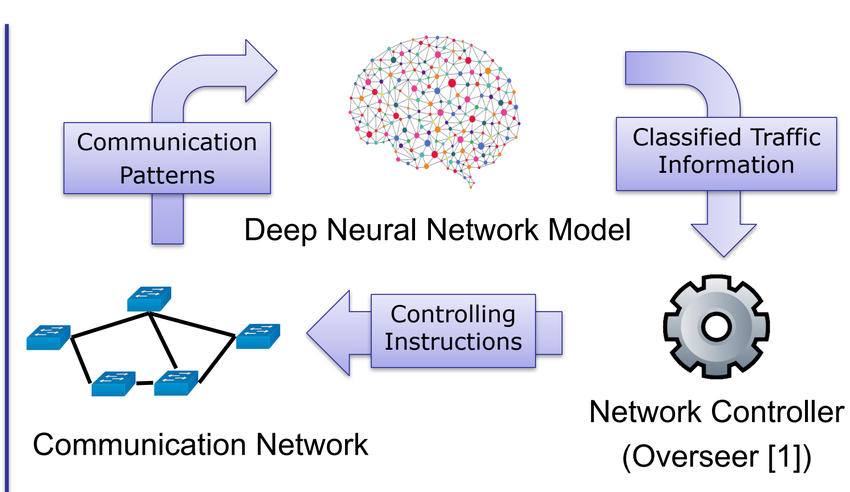
Communication network performance is essential, especially in this IoT era. Since network communications are diverse so as network path properties, they can be managed to maximize the efficiency. Communication enhancing cycle is a process that allows the network to self-optimize over time. The classification model is the missing piece in this puzzle.



Diverse communication patterns and path properties



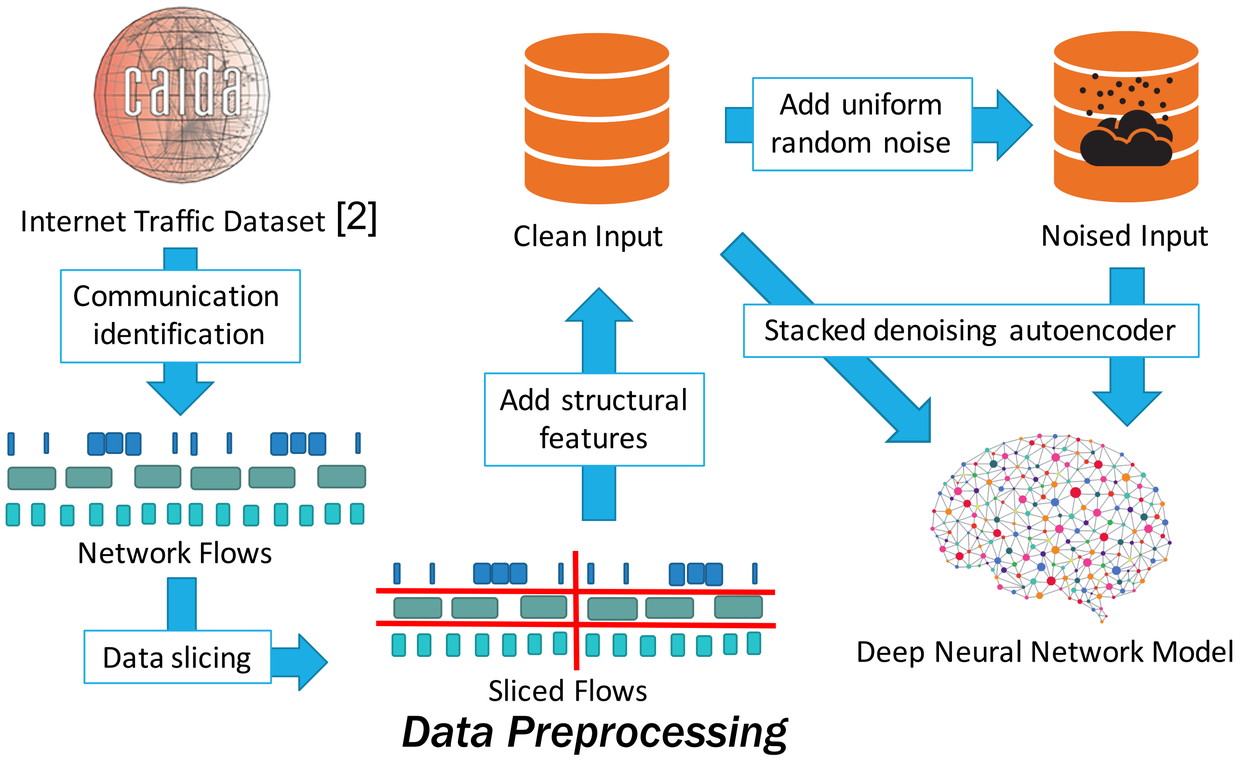
Managed communications



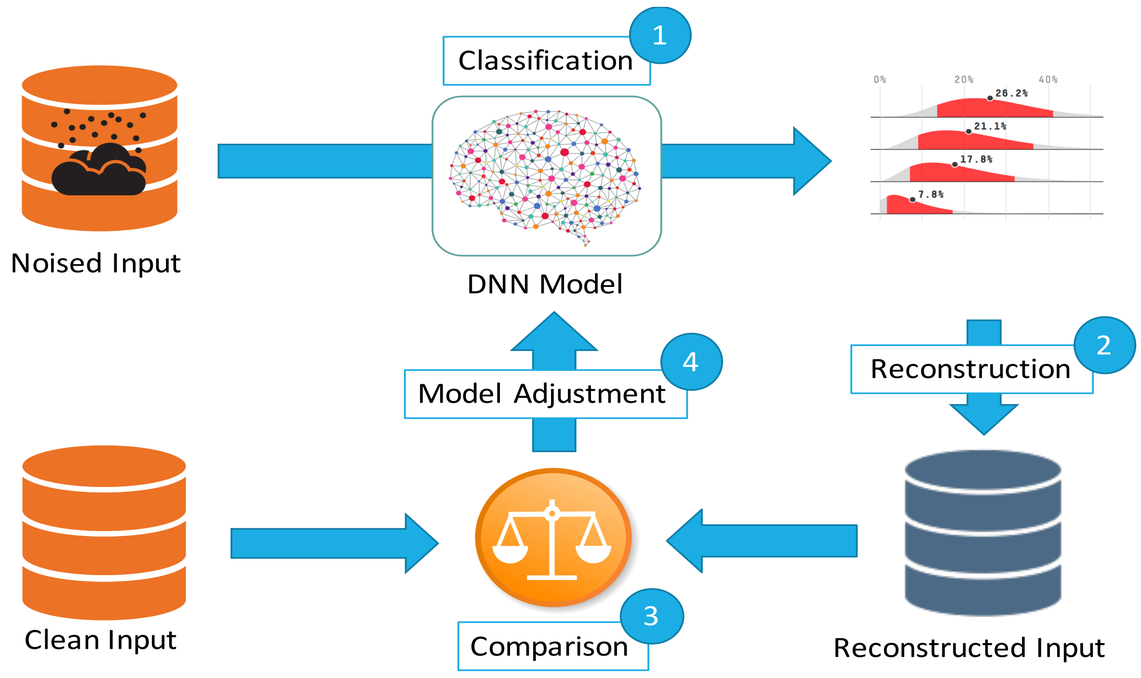
Communication Enhancing Cycle

Model Construction with Deep Learning

Deep learning is a powerful, proven, technique for recognizing complicated patterns. The model is developed using the (unsupervised) stacked denoising autoencoder algorithm. CAIDA Internet traffic, a large-scale raw Internet traffic dataset, is used to train the model [2]. The following figures illustrate model construction process and training process respectively.



Data Preprocessing



Training (Stacked Denoising Autoencoder)

Resulting Model

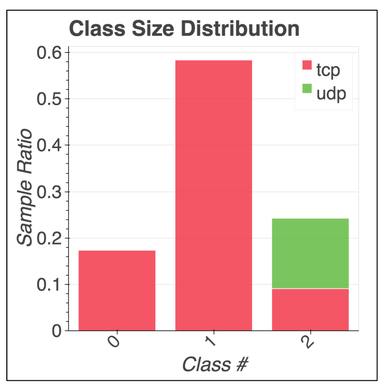
The resulting model classifies network communication into 3 classes. The classification is interpreted by observing distributions of several properties of each class.

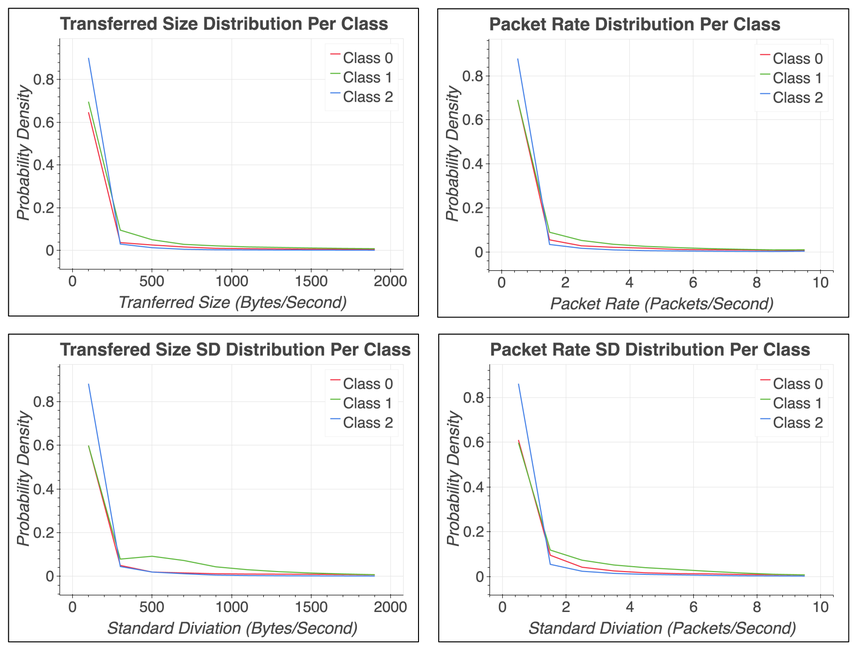
<u>Class 0:</u> Standard frequency communication with relatively low packet size (latency-bound pattern)

<u>Class 1:</u> Standard frequency communication with relatively high packet size (bandwidth-bound pattern)

<u>Class 2:</u> Low frequency communication

The results suggested that deep learning is a practical approach for network traffic identification and categorization. We are actively working on improving the quality of the model as well as deploying the model in a communication enhancing cycle.





Class Properties Distributions

[1] P. U-chupala et al., “Application-Oriented Bandwidth and Latency Aware Routing with OpenFlow Network,” in The 6th IEEE International Conference on Cloud Computing Technology and Science (CloudCom), 2014.
[2] CAIDA: the Cooperative Association for Internet Data Analysis. http://www.caida.org/