README

This data set can be divided in two components: (i) the data sets used to train the delay/jitter RoutNet models (‘nsfnet.zip’, ‘geant2.zip’, ‘gbn.zip’), and (ii) the delay/jitter RouteNet models already trained that we used for the evaluation in the paper (‘trained_model_delay.zip’, ‘trained_model_jitter.zip’).

Description of the training data sets:
- The root directory contains a *.ned file that describes the network topology. We consider that all the links in the topology have a capacity of 10 kbps.

- The /routing* directory contains the different routing schemes simulated. Each routing file includes a matrix that describes the forwarding of each node: matrix(src node , dst node) = output port of the src node to forward the traffic towards the dst node.

- The /delays* directory contains the results of the simulation. The file names use the following structure : dGlobal_0_<lambda>_<routing_file>.txt, where <lambda> is the traffic intensity and <routing_file> indicates the routing used for the simulation. Each line in these dGlobal_0_*.txt files corresponds to a simulation with different traffic matrices probabilistically generated from a given traffic intensity (lambda). We describe below the data structure in each line (i.e., each simulation). Note that in a topology with ‘n’ nodes, nodes are enumerated in the range [0, n-1].

1. Bandwidth (in kbps) generated for each source-destination pair in the network. This is a flattened vector with nxn fields, where the index for a particular src-dst pair is:
   $$index = src_{node} * n + dst_{node}$$
   For example, the position of the bw generated for the pair 0-7 in a network with 10 nodes would be index = 0*10+7 = 7.
2. Average delay over the packets transmitted for each source-destination pair in the network. This contains nxn values. This vector and the subsequent have the same indexes for nodes than the previous vector.
3. Percentile 10 over the packets transmitted for each source-destination pair. This contains nxn values.
4. Percentile 20 over the packets transmitted for each source-destination pair. This contains nxn values.
5. Percentile 50 over the packets transmitted for each source-destination pair. This contains nxn values.
6. Percentile 80 over the packets transmitted for each source-destination pair. This contains nxn values.
7. Percentile 90 over the packets transmitted for each source-destination pair. This contains nxn values.
8. Variance of the delay (jitter) over the packets transmitted for each source-destination pair. This contains nxn values.
9. Absolute number of packet losses in each source-destination pair. This contains nxn values
10. Total number of packets lost in the network. This is the sum of all the losses in the previous vector.

In order to use the delay/jitter RouteNet models already trained, you can find more information at https://github.com/knowledgedefinednetworking/net2vec/tree/master/routenet.

If you have any doubts, do not hesitate to contact us by sending an email to kdn-contactus <at> knowledgedefinednetworking.org. Also, if you would like to be notified with new data set releases or discuss anything related to the data sets, you can also subscribe to the mailing list kdn-users<at> knowledgedefinednetworking.org (Link: https://mail.knowledgedefinednetworking.org/cgi-bin/mailman/listinfo/kdn-users).