GLOBECOM 2022 - Rio de Janeiro (Brazil)

NeCSTGen: An approach for realistic network traffic generation using Deep Learning.

Authors :

- Fabien MESLET-MILLET (IRIT/Toulouse INP)
- Sandrine MOUYSSET (IRIT/Toulouse UPS)
- Emmanuel CHAPUT (IRIT/Toulouse INP)

Table of contents

Presentation of the problem

- Goals
- Formalization

Presentation of the problem

Goals

- Goals
  - Perform network traffic generation, without payload, by using Deep Learning methods.
- Interests
  - IoT (Internet of Things)
  - Protocol agnostic
  - Ground truth
- Issues
  - Generate dynamics and variables together
  - Dealing with long sequences
  - Dealing with different protocols
  - Multi-level (packet, flow, flow aggregate, …)

Presentation of the problem

Formalization

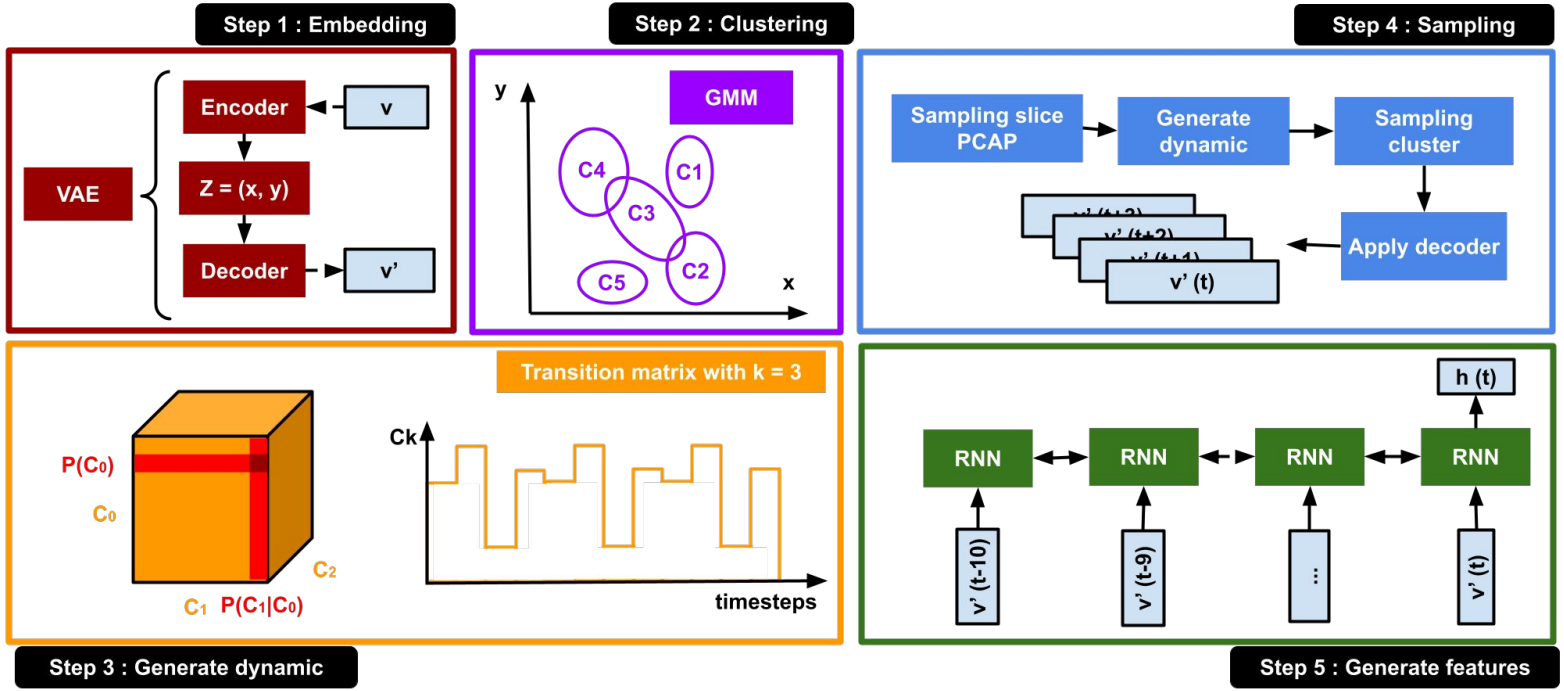
- Datasets
  - IoT
    - LoraWan
    - Google Home
  - LAN
    - DARPA
- Generation
  - Protocol diversity
  - Application diversity
  - Dynamic diversity
- Header
  - IP
  - TCP / UDP
- Payload
  - SMTP
  - HTTP
  - SNMP
  - …etc

NeCSTGen architecture

- Overview
- Step 1: Projection
- Step 2: Clustering
- Step 3: Generate dynamic
- Step 4: Sampling
- Step 5: Generate features
- Multi-levels

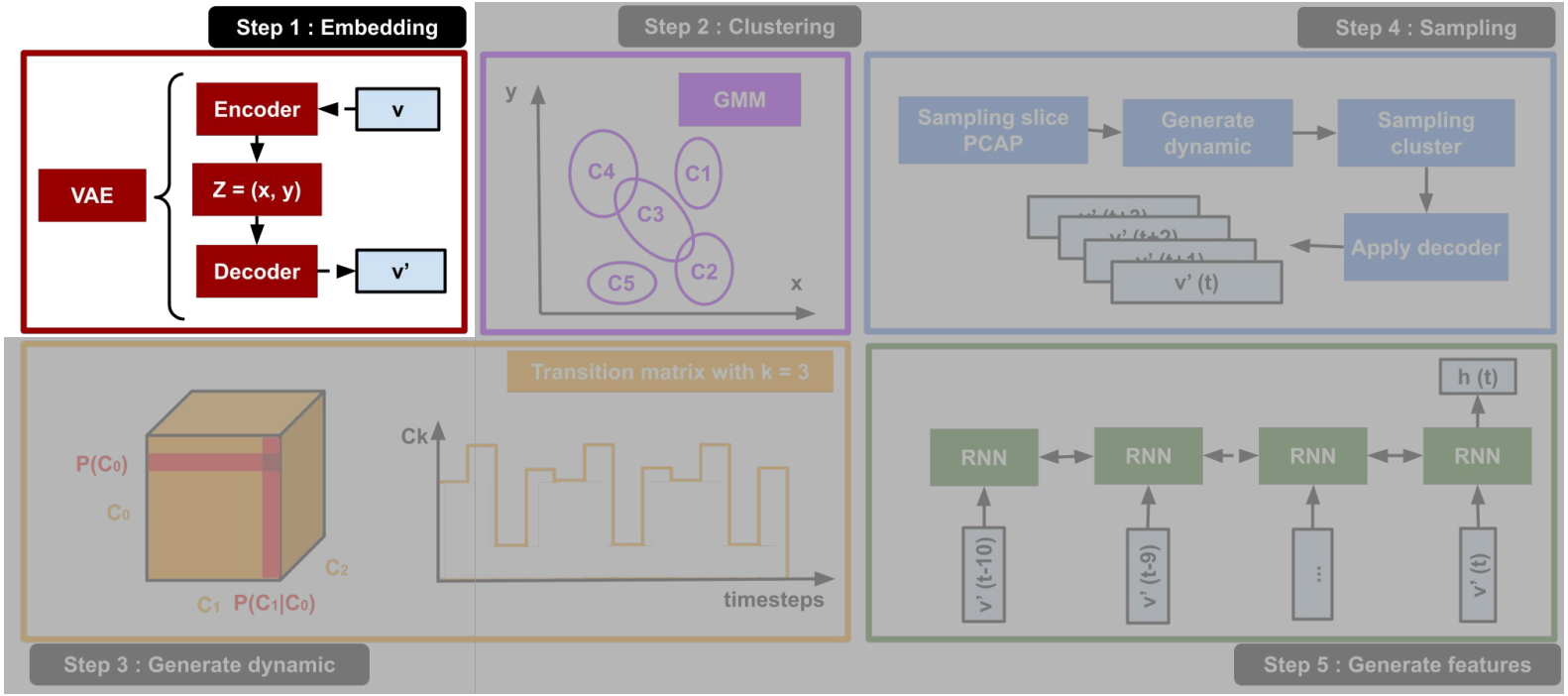
NeCSTGen architecture

Overview



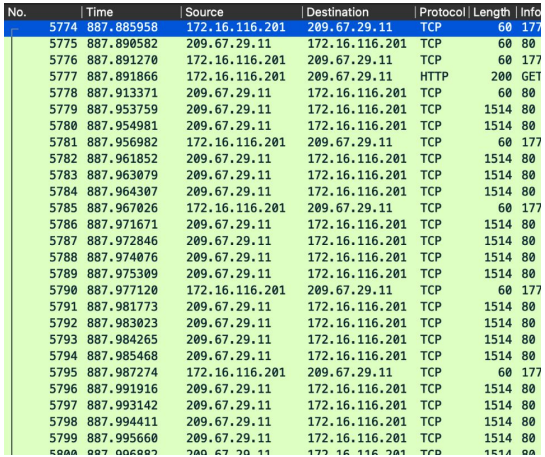
NeCSTGen architecture

Step 1: Embedding (1/3)



NeCSTGen architecture

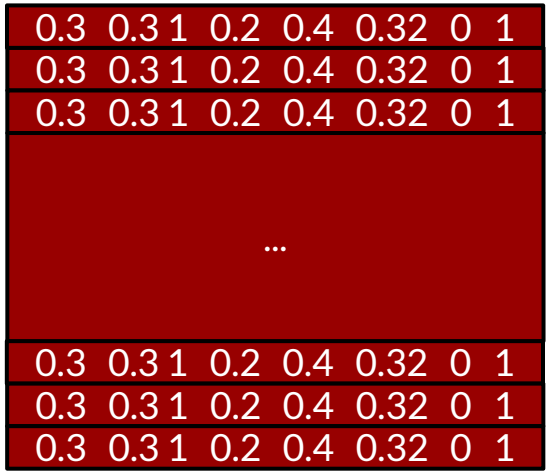
Step 1: Embedding (2/3)



No.	Time	Source	Destination	Protocol	Length	Info
5774	887.885958	172.16.116.201	209.67.29.11	TCP	60	177
5775	887.890582	209.67.29.11	172.16.116.201	TCP	60	80
5776	887.891270	172.16.116.201	209.67.29.11	TCP	60	177
5777	887.891866	172.16.116.201	209.67.29.11	HTTP	200	GET
5778	887.913371	209.67.29.11	172.16.116.201	TCP	60	80
5779	887.953759	209.67.29.11	172.16.116.201	TCP	1514	80
5780	887.954981	209.67.29.11	172.16.116.201	TCP	1514	80
5781	887.956982	172.16.116.201	209.67.29.11	TCP	60	177
5782	887.961852	209.67.29.11	172.16.116.201	TCP	1514	80
5783	887.963079	209.67.29.11	172.16.116.201	TCP	1514	80
5784	887.964307	209.67.29.11	172.16.116.201	TCP	1514	80
5785	887.967026	172.16.116.201	209.67.29.11	TCP	60	177
5786	887.971671	209.67.29.11	172.16.116.201	TCP	1514	80
5787	887.972846	209.67.29.11	172.16.116.201	TCP	1514	80
5788	887.974076	209.67.29.11	172.16.116.201	TCP	1514	80
5789	887.975309	209.67.29.11	172.16.116.201	TCP	1514	80
5790	887.977120	172.16.116.201	209.67.29.11	TCP	60	177
5791	887.981773	209.67.29.11	172.16.116.201	TCP	1514	80
5792	887.983023	209.67.29.11	172.16.116.201	TCP	1514	80
5793	887.984265	209.67.29.11	172.16.116.201	TCP	1514	80
5794	887.985468	209.67.29.11	172.16.116.201	TCP	1514	80
5795	887.987274	172.16.116.201	209.67.29.11	TCP	60	177
5796	887.991916	209.67.29.11	172.16.116.201	TCP	1514	80
5797	887.993142	209.67.29.11	172.16.116.201	TCP	1514	80
5798	887.994411	209.67.29.11	172.16.116.201	TCP	1514	80
5799	887.995660	209.67.29.11	172.16.116.201	TCP	1514	80
5800	887.996982	209.67.29.11	172.16.116.201	TCP	1514	80

PCAP including an aggregate of HTTP flows

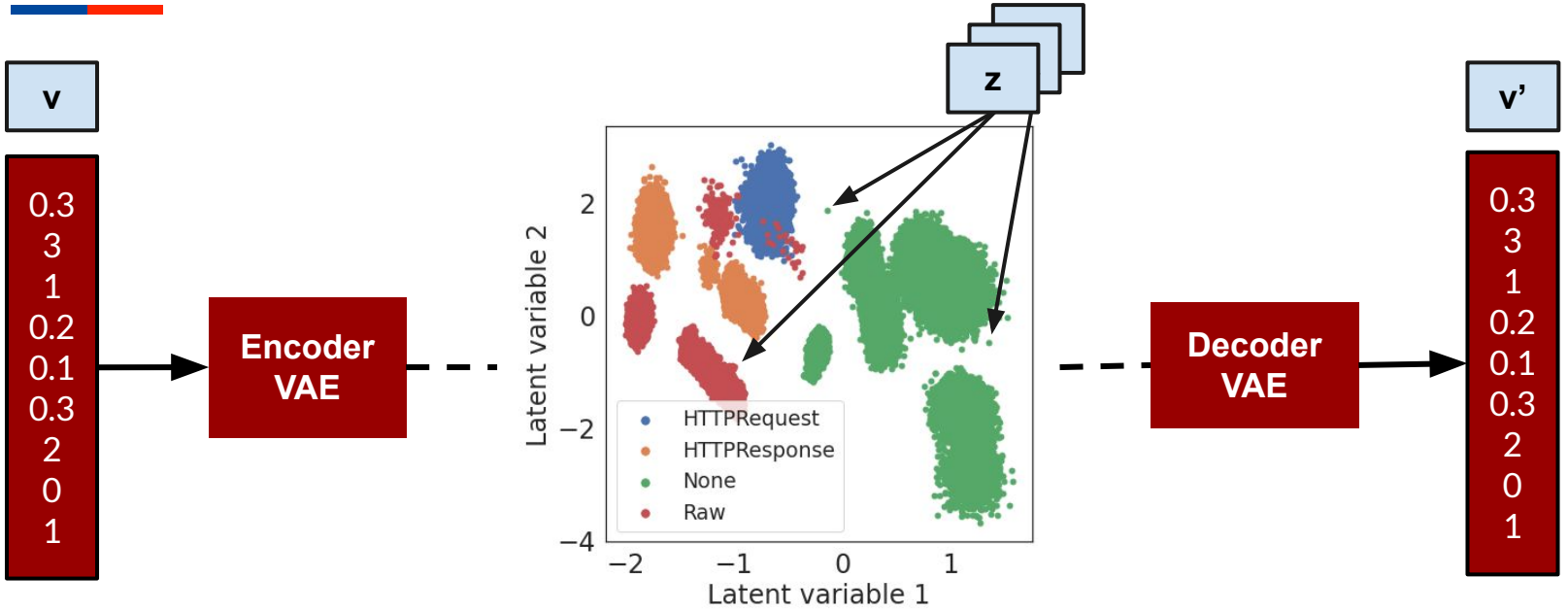
Features extraction for each packet



Features vector generated for **1 packet** (level 2/3/4 header, rate in bytes/s...etc)

NeCSTGen architecture

Step 1: Embedding (3/3)



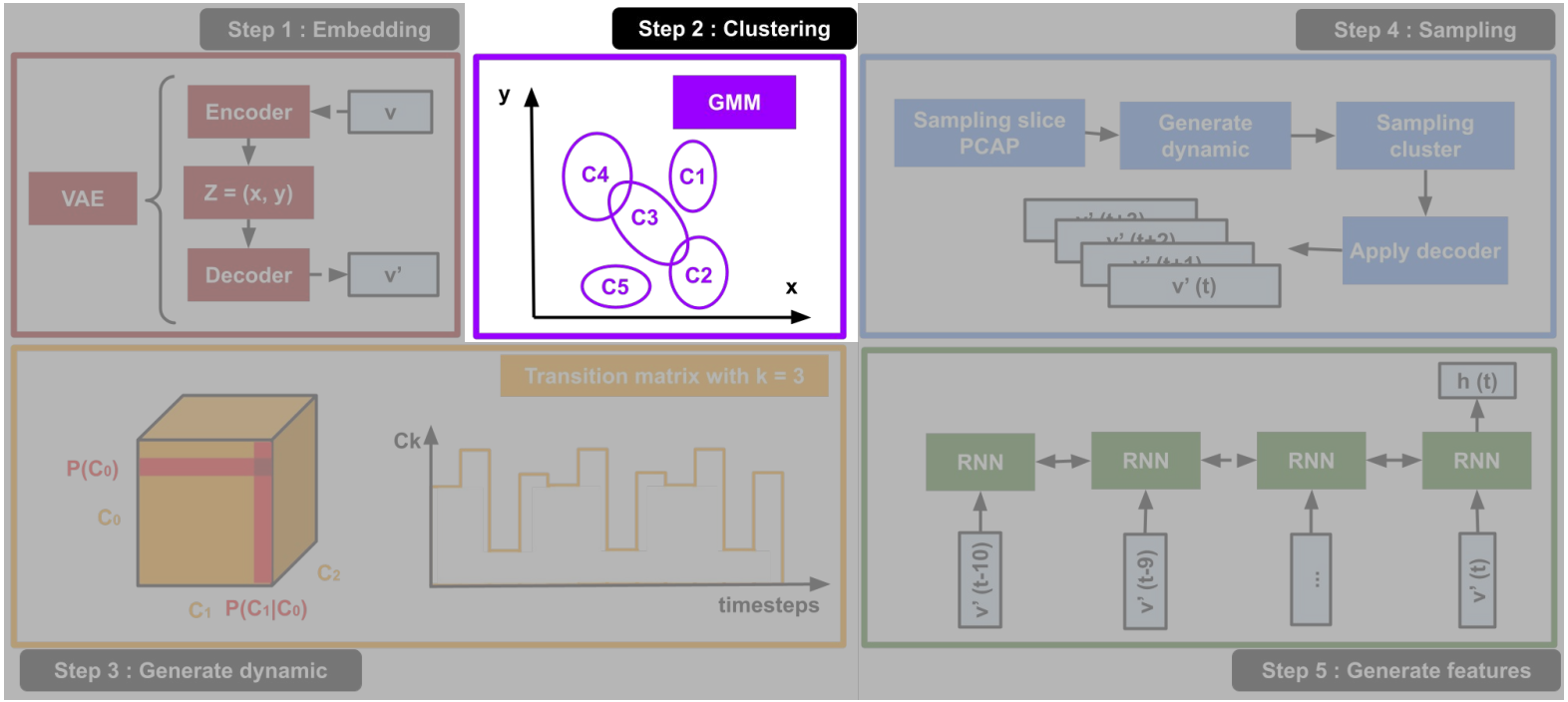
Vector of features for **1 packet** (header of level 2/3/4, rate in bytes/s...etc)

Latent space of the packet projection of an HTTP packet aggregate

Vector of features generated for **1 packet** (header of level 2/3/4, rate in bytes/s...etc)

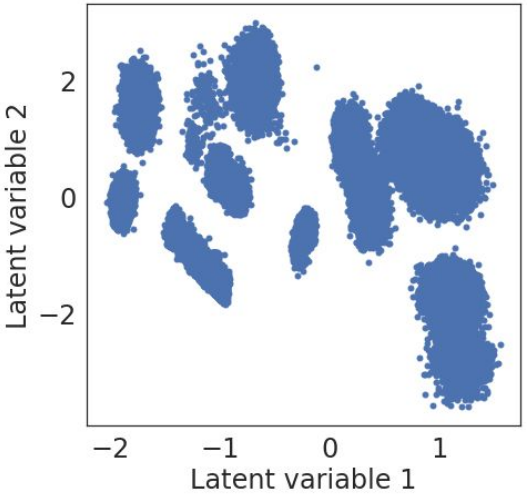
NeCSTGen architecture

Step 2: Clustering (1/3)

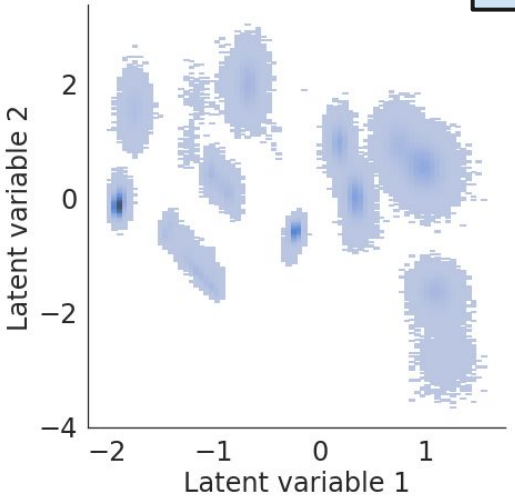


NeCSTGen architecture

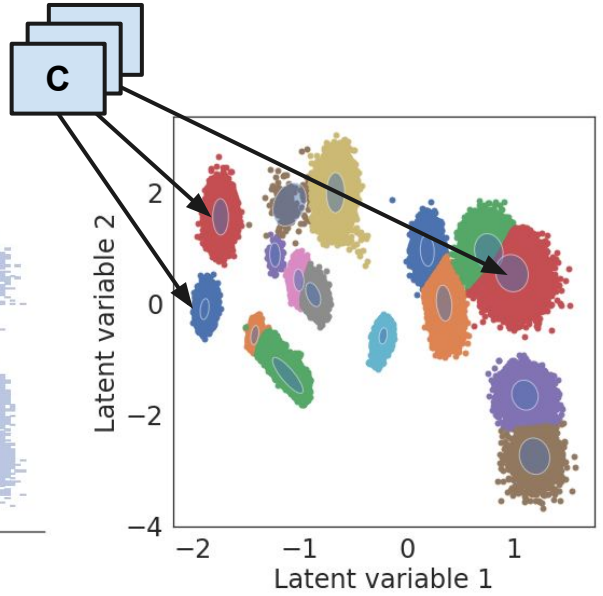
Step 2: Clustering HTTP (2/3)



Projection of the features of each packet by the VAE.



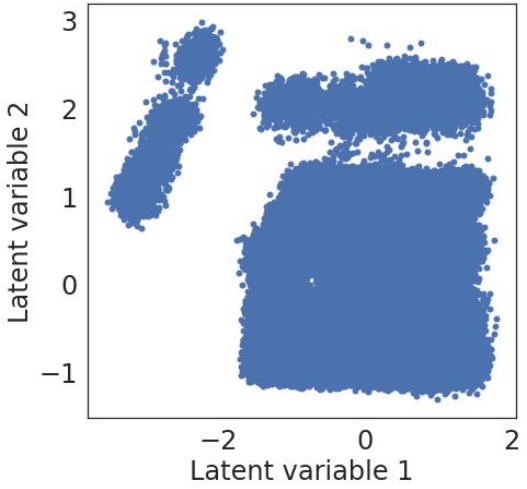
Distribution of the projection of the features of each packet by the VAE.



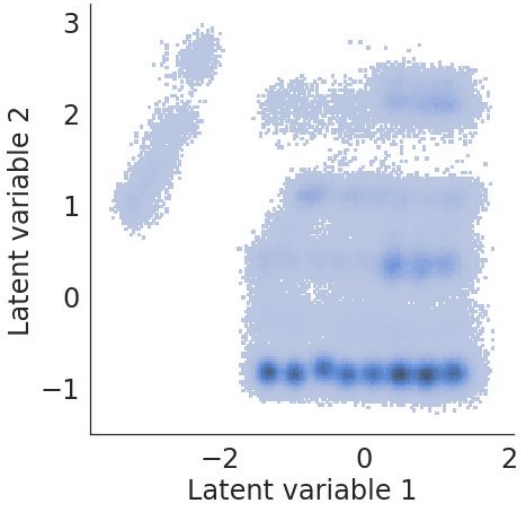
Clustering with GMM of the projection of the features of each packet by the VAE.

NeCSTGen architecture

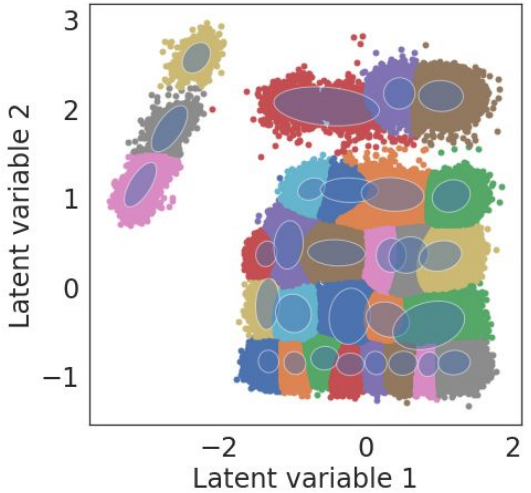
Step 2: Clustering LoraWAN (fport: 1) (3/3)



Projection of the features of each packet by the VAE.



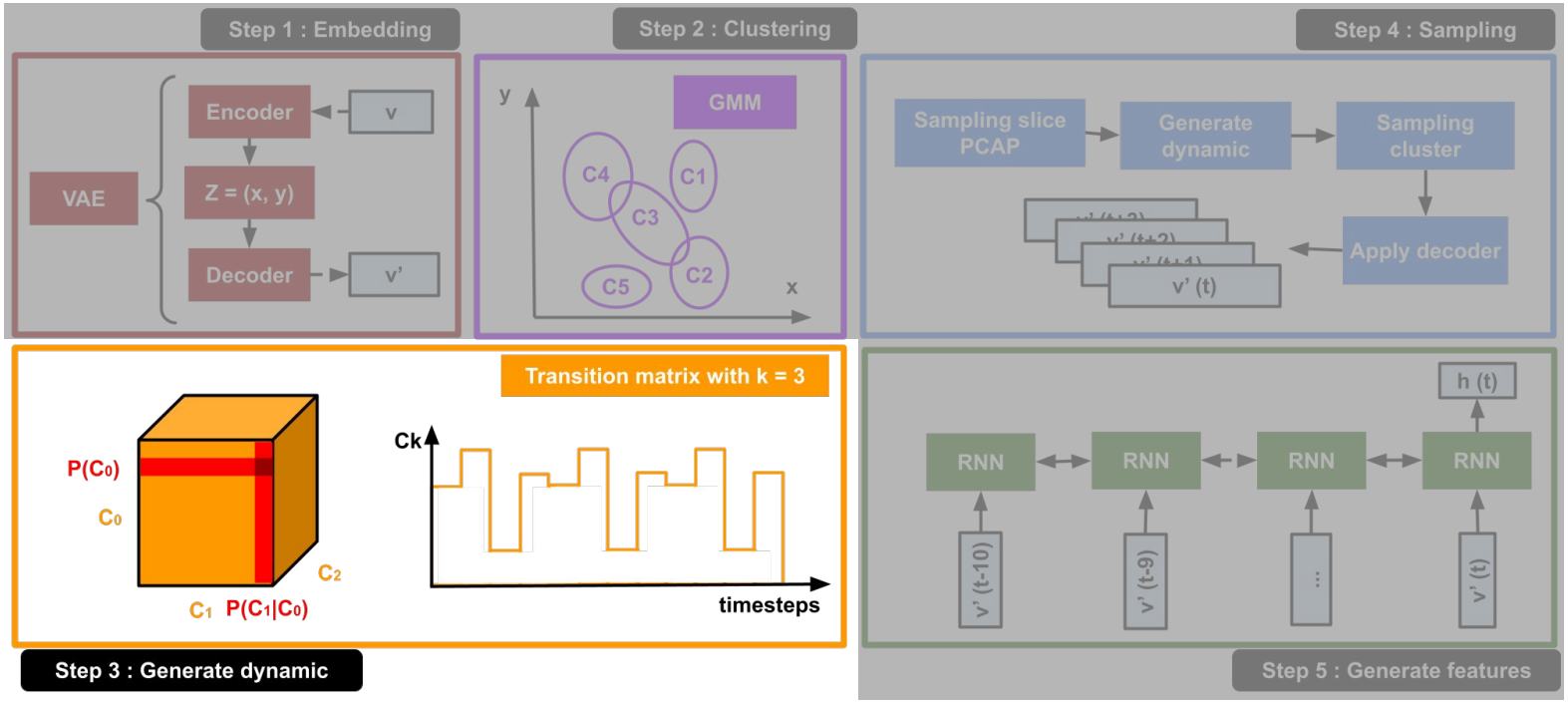
Distribution of the projection of the features of each packet by the VAE.



Clustering with GMM of the projection of the features of each packet by the VAE.

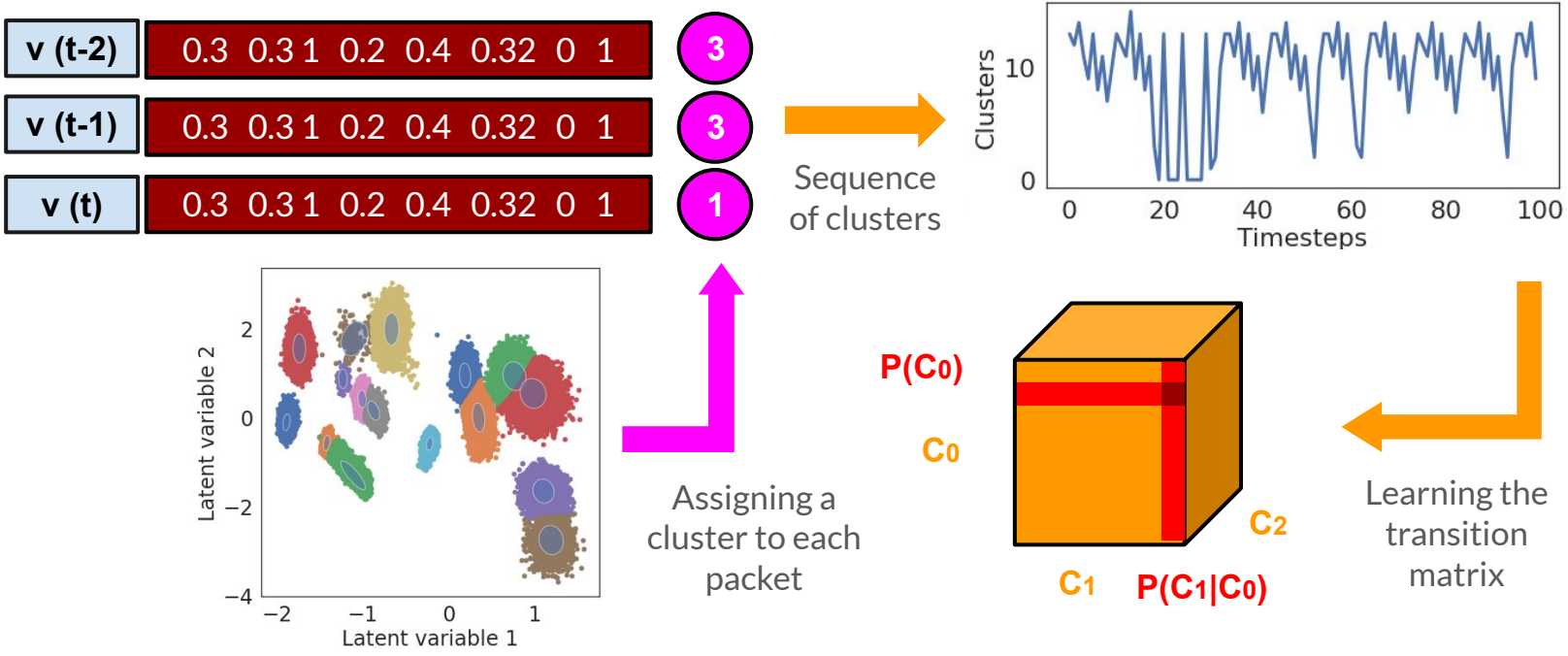
NeCSTGen architecture

Step 3: Generate dynamic (1/2)



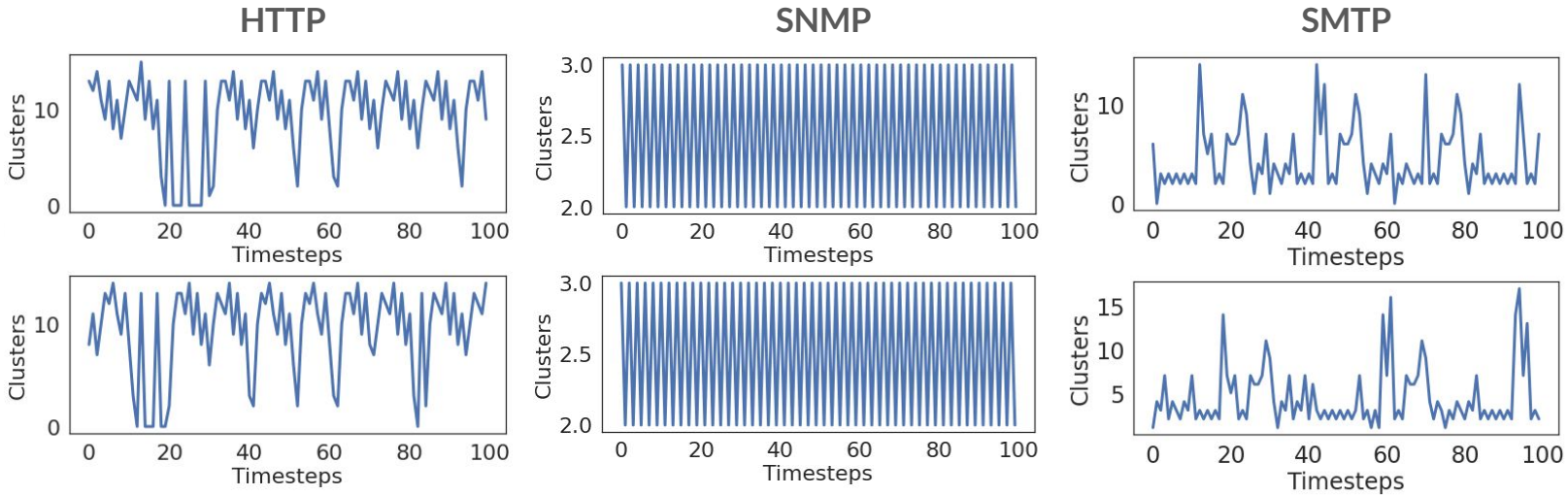
NeCSTGen architecture

Step 3: Generate dynamic (2/3)



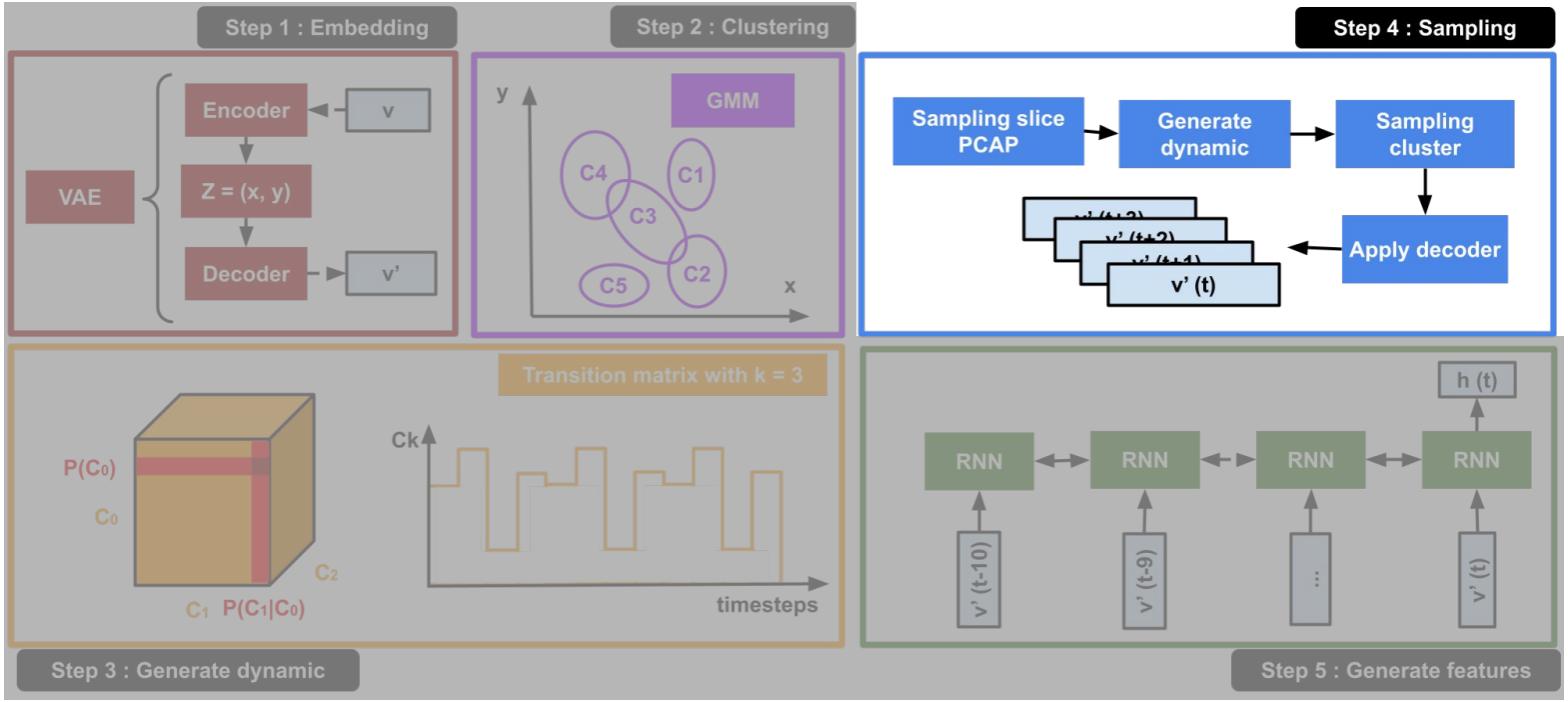
NeCSTGen architecture

Step 3: Generate dynamic (3/3)



NeCSTGen architecture

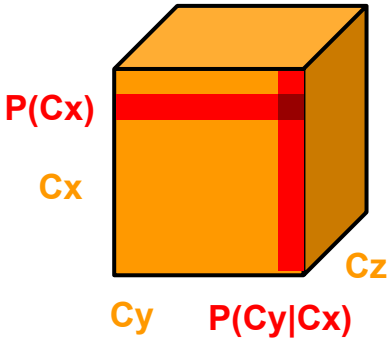
Step 4: Sampling (1/2)

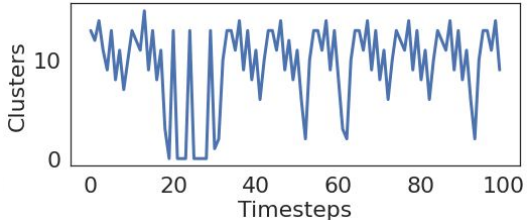


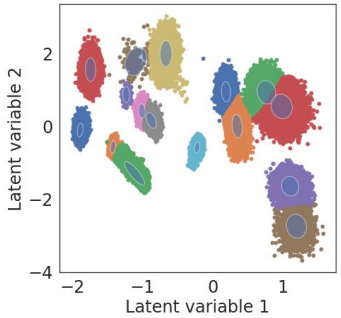
NeCSTGen architecture

Step 4: Sampling (2/2)

Sampling of the transition matrix







Sampling of the clusters

z' (t)

x
y

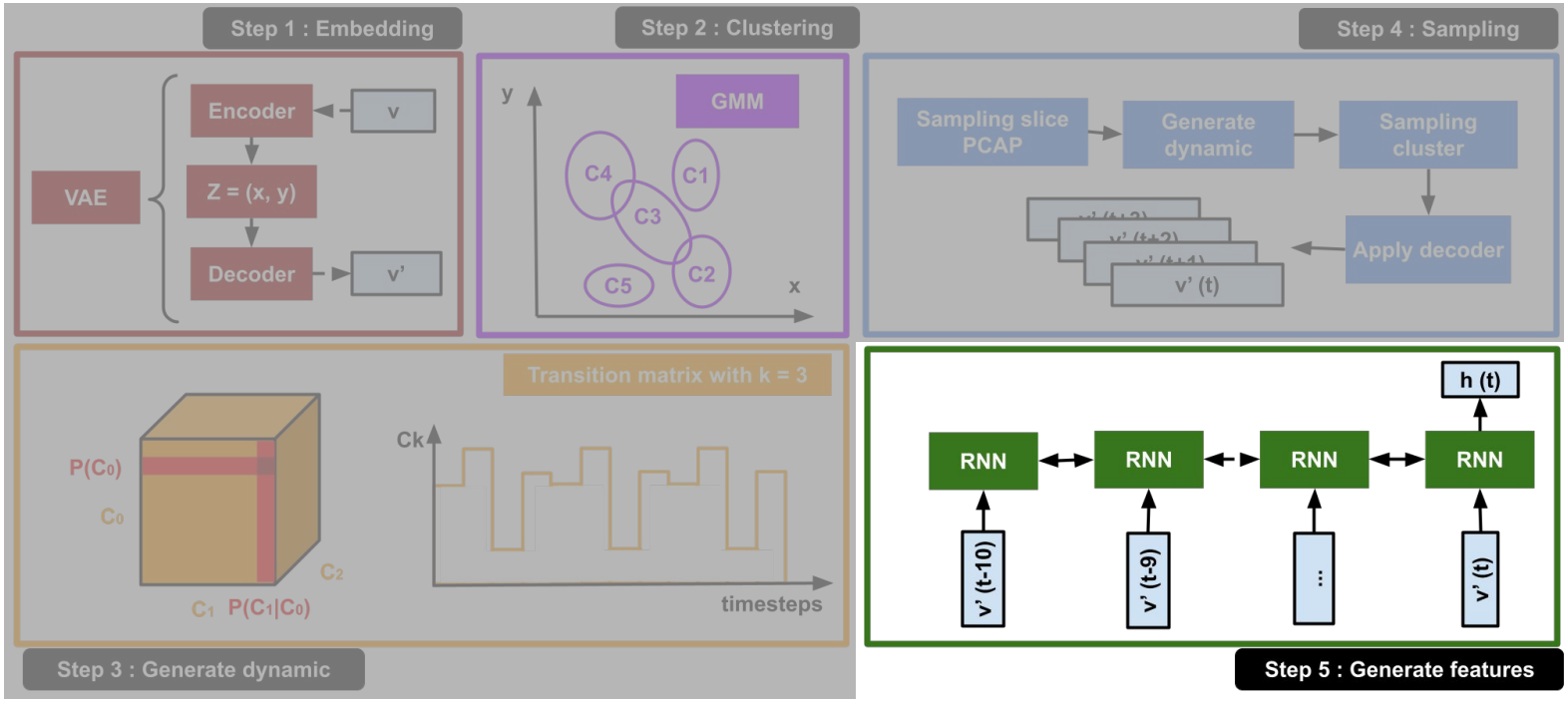
Decoder VAE

v' (t)

0.3
3
1
0.2
0.1
0.3
2
0
1

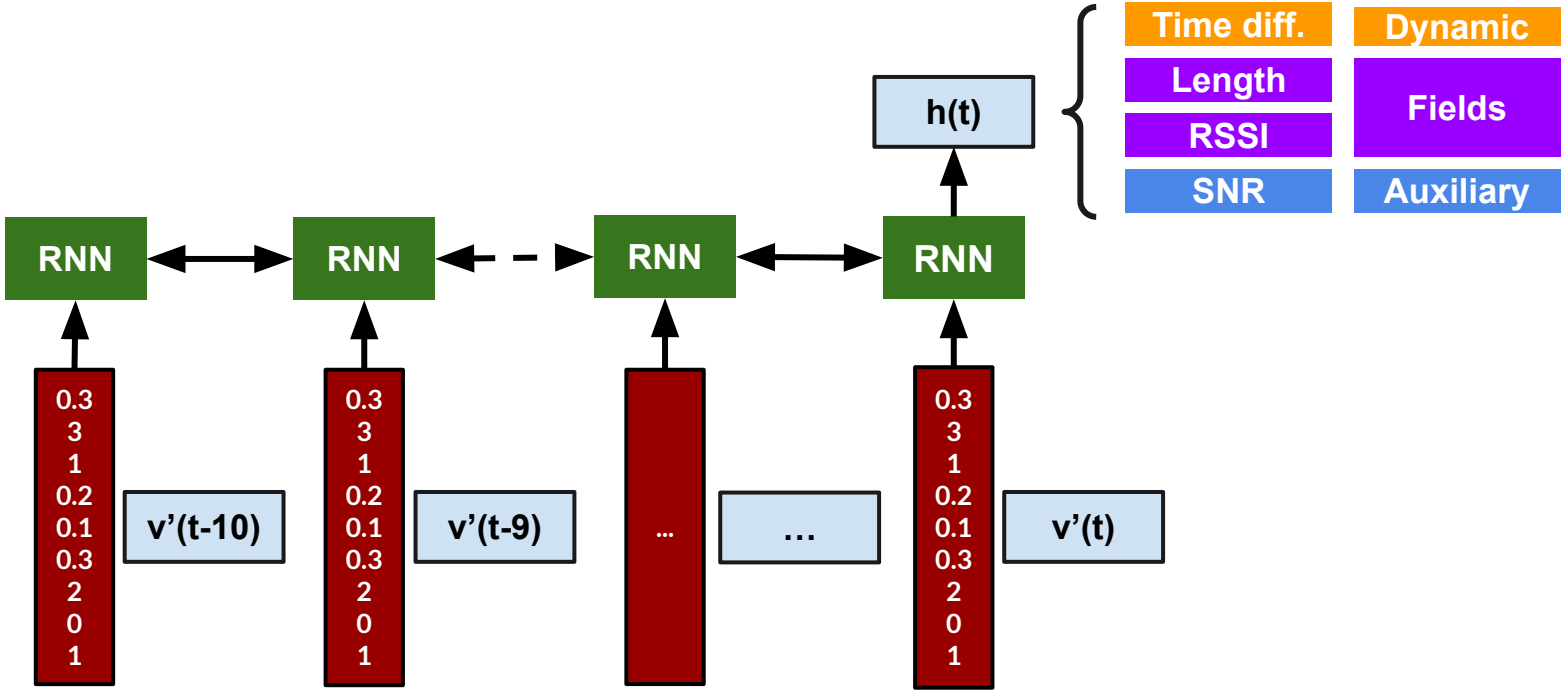
NeCSTGen architecture

Step 5: Generate features (1/2)



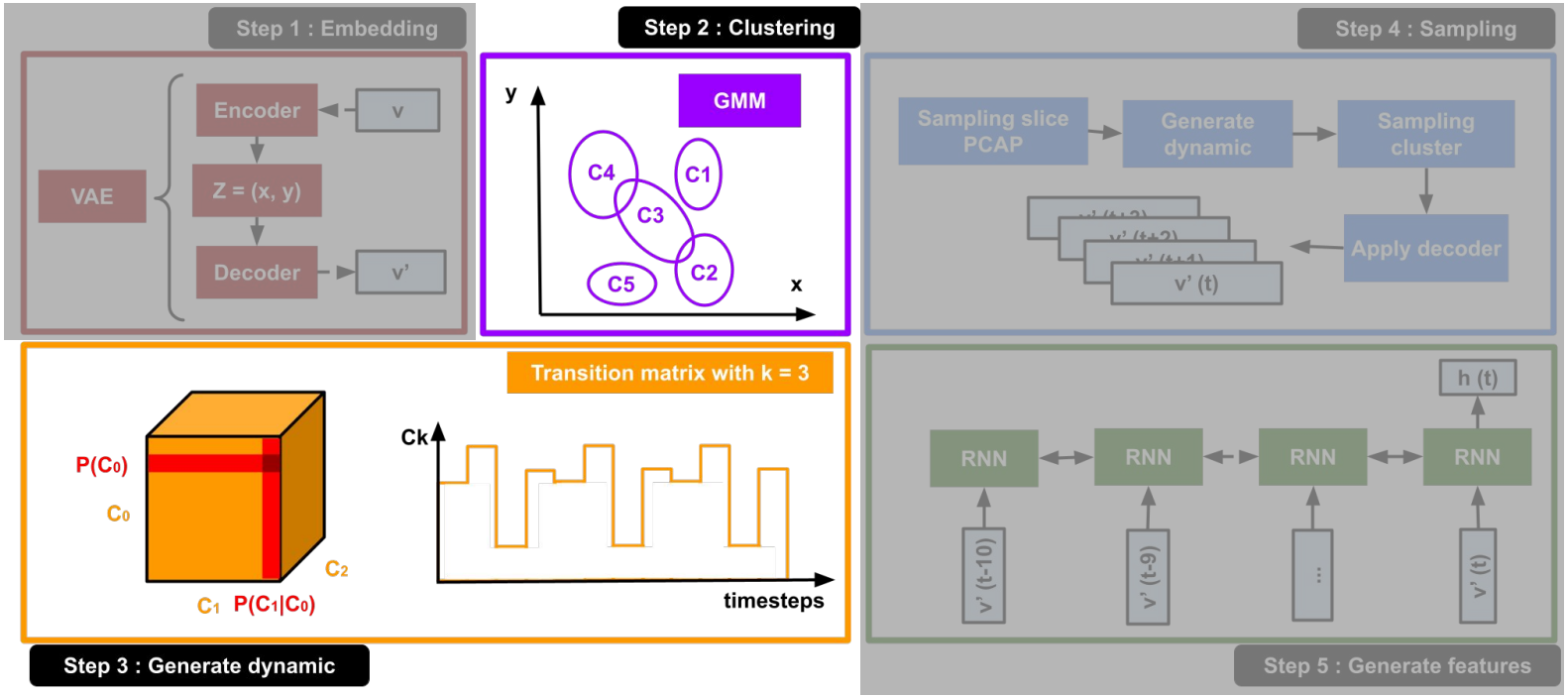
NeCSTGen architecture

Step 5: Generate features



NeCSTGen architecture

Multi-levels

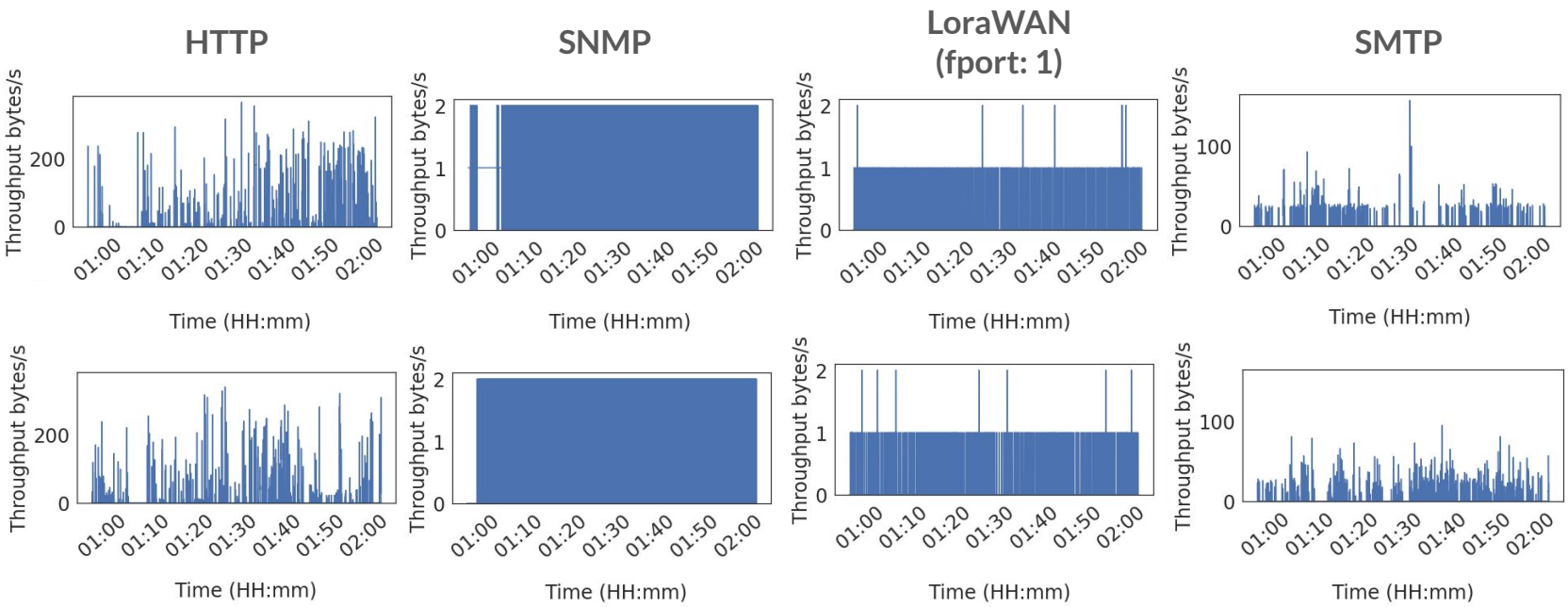


Results

- Packet dynamic
- Packet fields
- Flow and aggregate dynamics
- Scale characteristics
- QoS/QoE

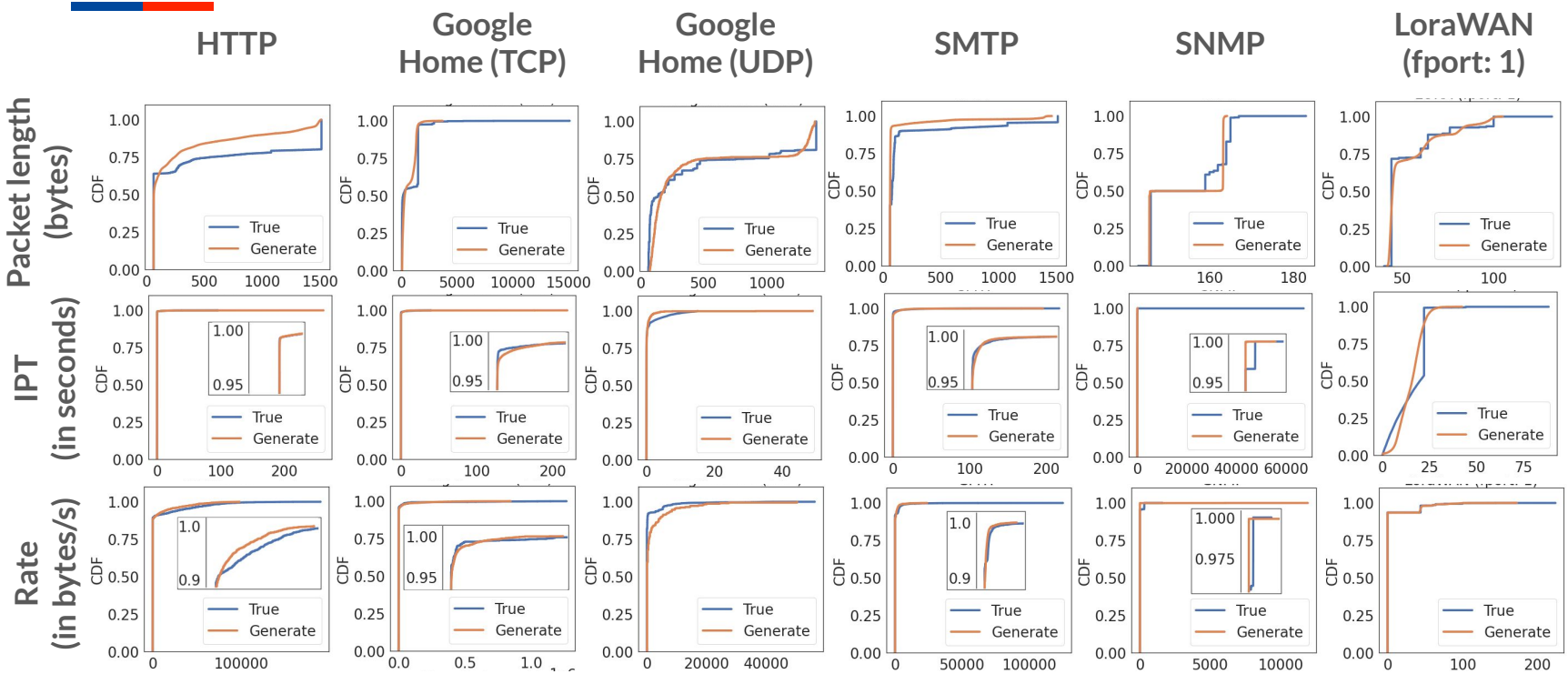
Results

Packet dynamic (1/2)



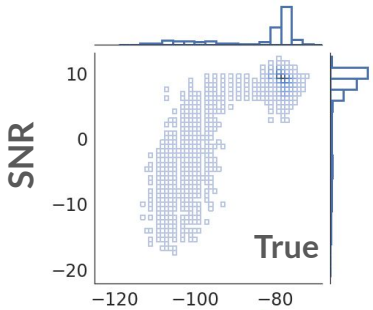
Results

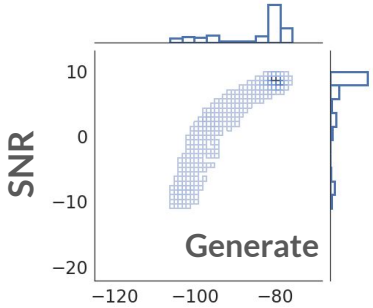
Packet dynamic (2/2)



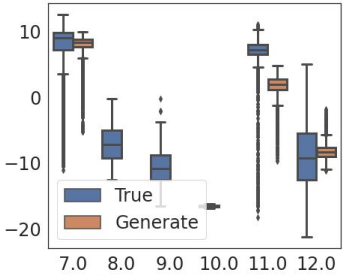
Results

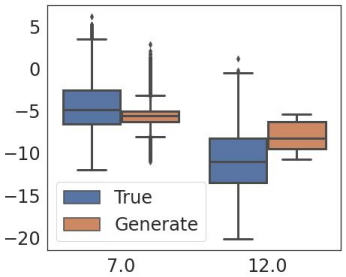
Packet fields



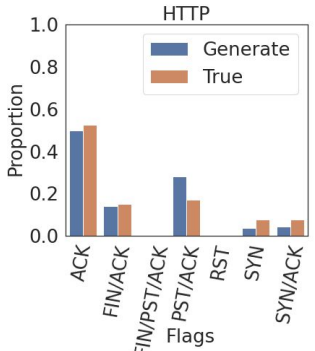


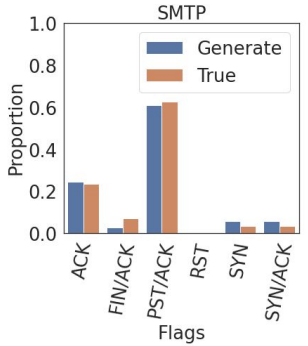
RSSI

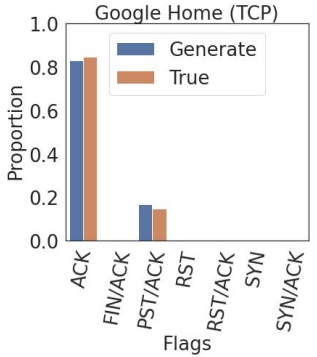




Spreading Factor

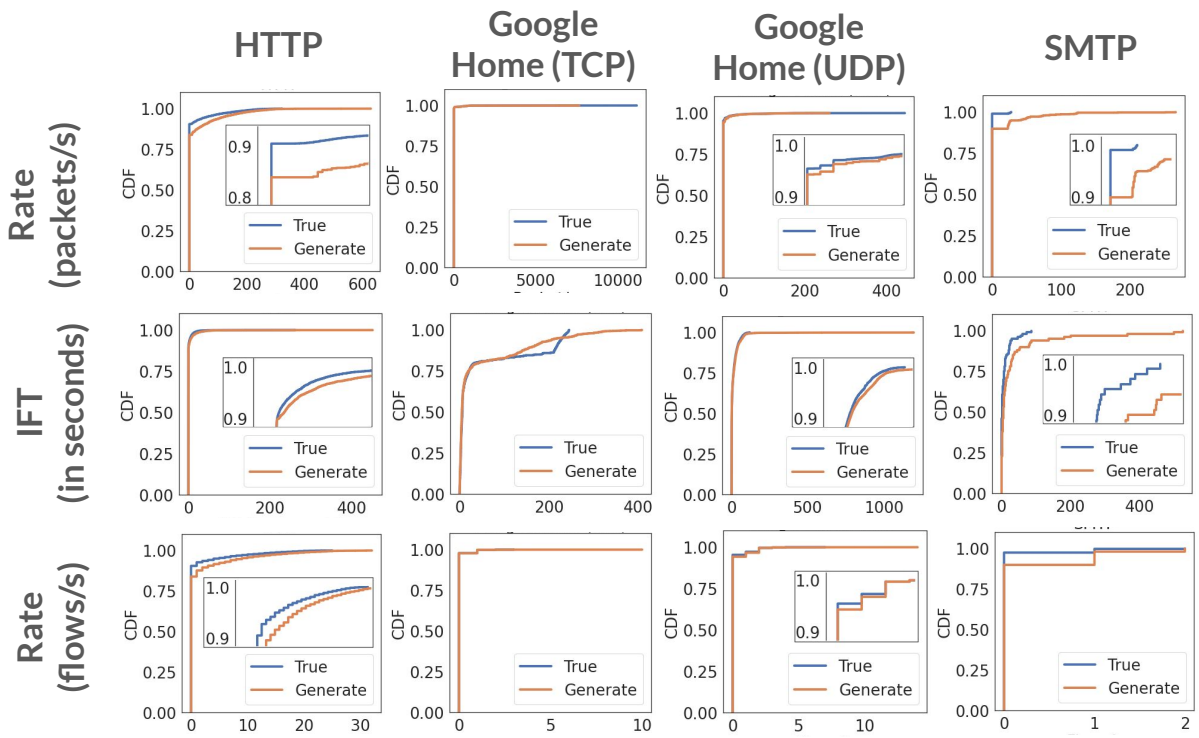






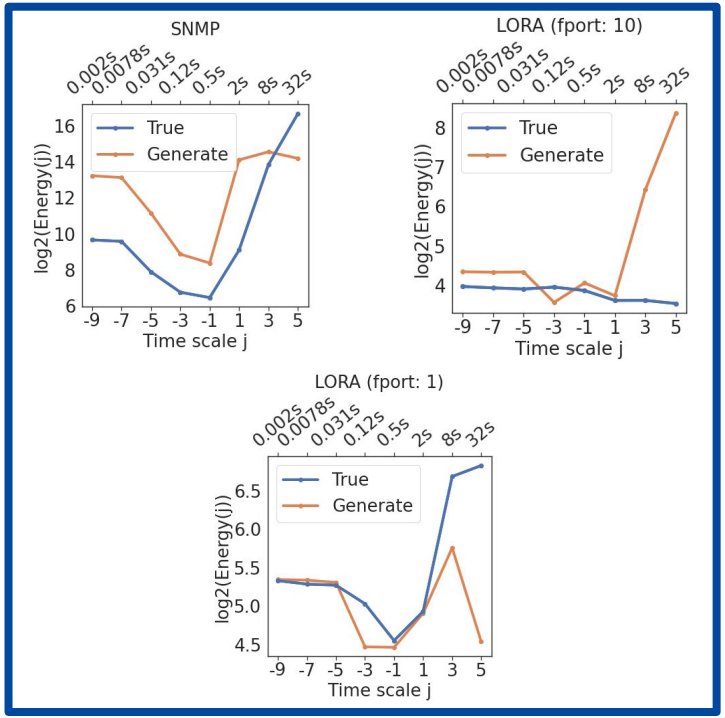
Results

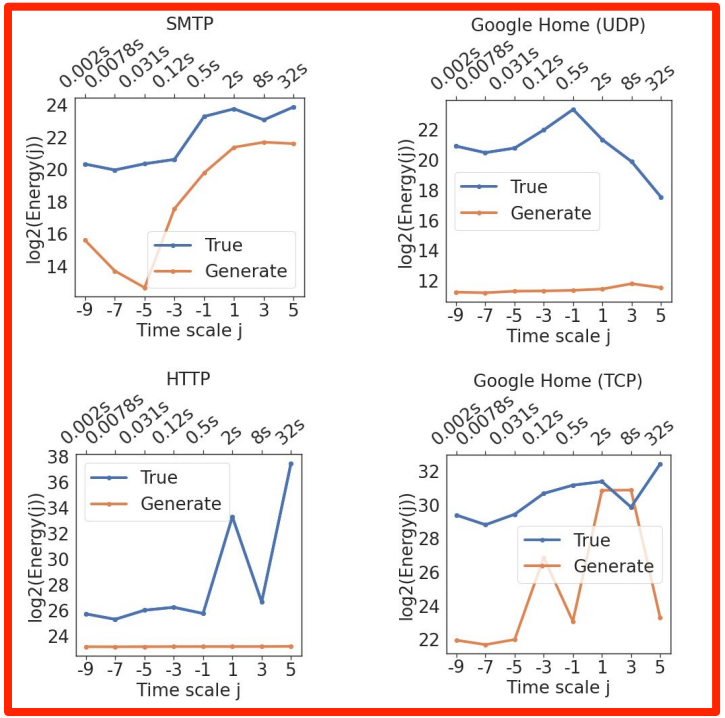
Flows and aggregate dynamic



Results

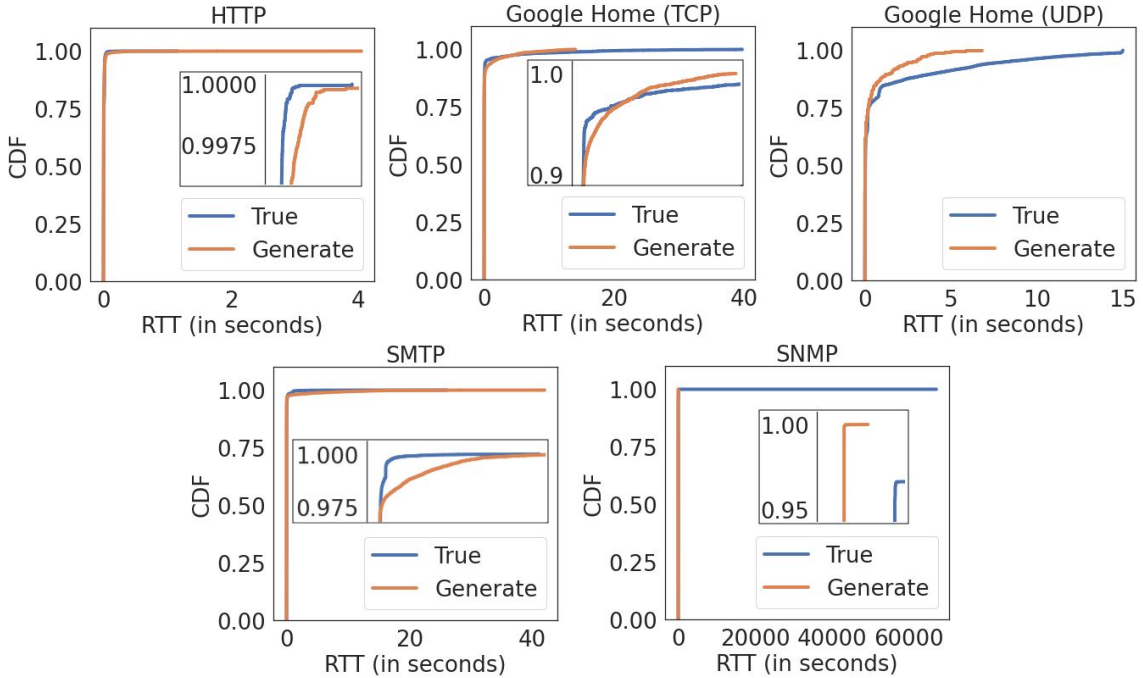
Scale characteristics





Results

QoS/QoE



Contributions

- Generation
  - Protocol agnostic
  - Ground truth
  - Generate dynamics and variables together
  - Multi-level (packet, flow, flow aggregate)
- Limits
  - Dimension of the transition matrix
  - Projection / Identification (with GMM)
- Paths of exploration
  - Dealing with rare events
  - Include variability of the traffic related to the human