Forecasting Model for Service Allocation Network Using Traffic Recognition

Ghazal Rahmanian[1] **(Presenter)**

Mohammad Yekta[1]

HadiShahriar Shahhoseini[1]

AmirHossein Jafari Pozveh [1,2]

[1]Iran University of Science and Technology (IUST)

[2]Iran Mobile Communications Company (MCI)

Introduction

01

Researcher's Background
IUST,ICT-SIS, SCaN

02

The problem statement & the team's contribution
Data preprocessing, ML algorithms

03

Results & evaluation

04

Question & Answer

IUST Research Centers

SuperComputing and Networking Research Lab (SCaN)

"IUST Research Center for ICT Strategic and International Studies" (ICT-SIS)

The Problem Statement

Description

Focusing on the intelligent application demand of networking management and computing resource management, the artificial intelligence technologies such as machine learning and big data include the possibilities of the softwarized approach in IMT-2020 (SDN/NFV) are applied to digital upgrade of the internet infrastructure. The one of the main issues in this area - is the services traffic allocation, taking into account the users dynamics. Here we propose the problem statement with the services traffic forecasting based on the changing user needs for services.
The suggestion of problem statement:

- Proposal with ML model for recognizing the user demands based on the traffic services allocation;
- Proposal with ML model for traffic forecasting, taking into account traffic types and user demands (in order to future service migration).

Dataset:
A 20GB zipped pcap format data including several different captures

Preprocess Target:
Deriving a matrix of network features in different types of sessions

Main Target:
Developing a ML model for network traffic prediction

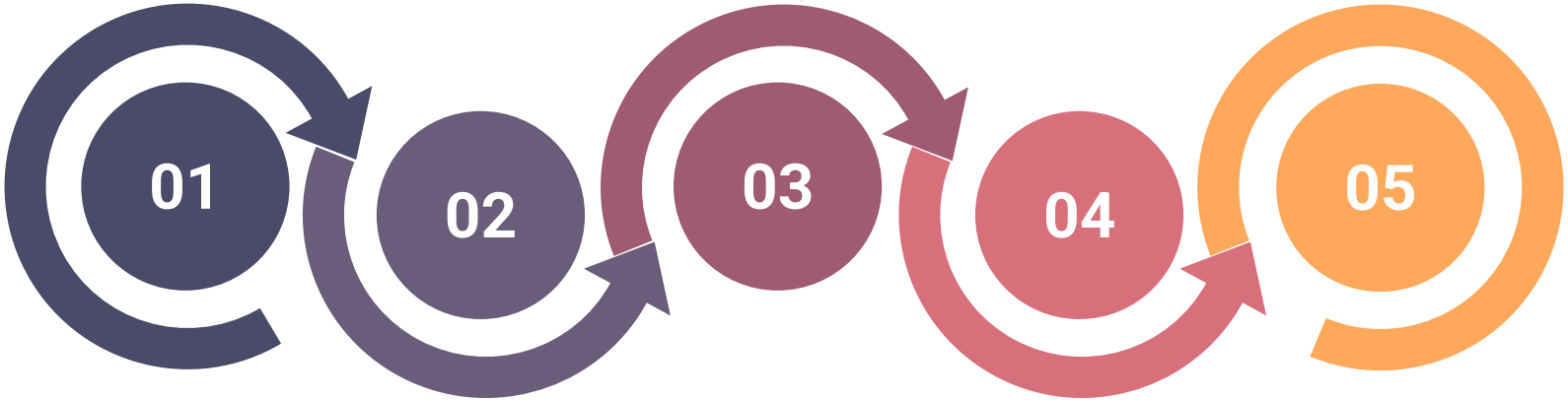
Research Procedure

Data Gathering

Provided by the challenge composers

Implementing ML Algorithms (SVR & NAR)

Comparison and Conclusion

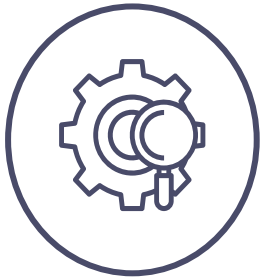


Data Preprocess

Evaluation and Results

Data Preprocessing

Conversion and Labeling



Converting the pcap file to CSV and labeling the dataset

Extraction

Extracting number of completed communications

Traffic Features Matrix

Developing the matrix of different connections among servers, clients & applications

Data Preprocessing (Number of Connections)

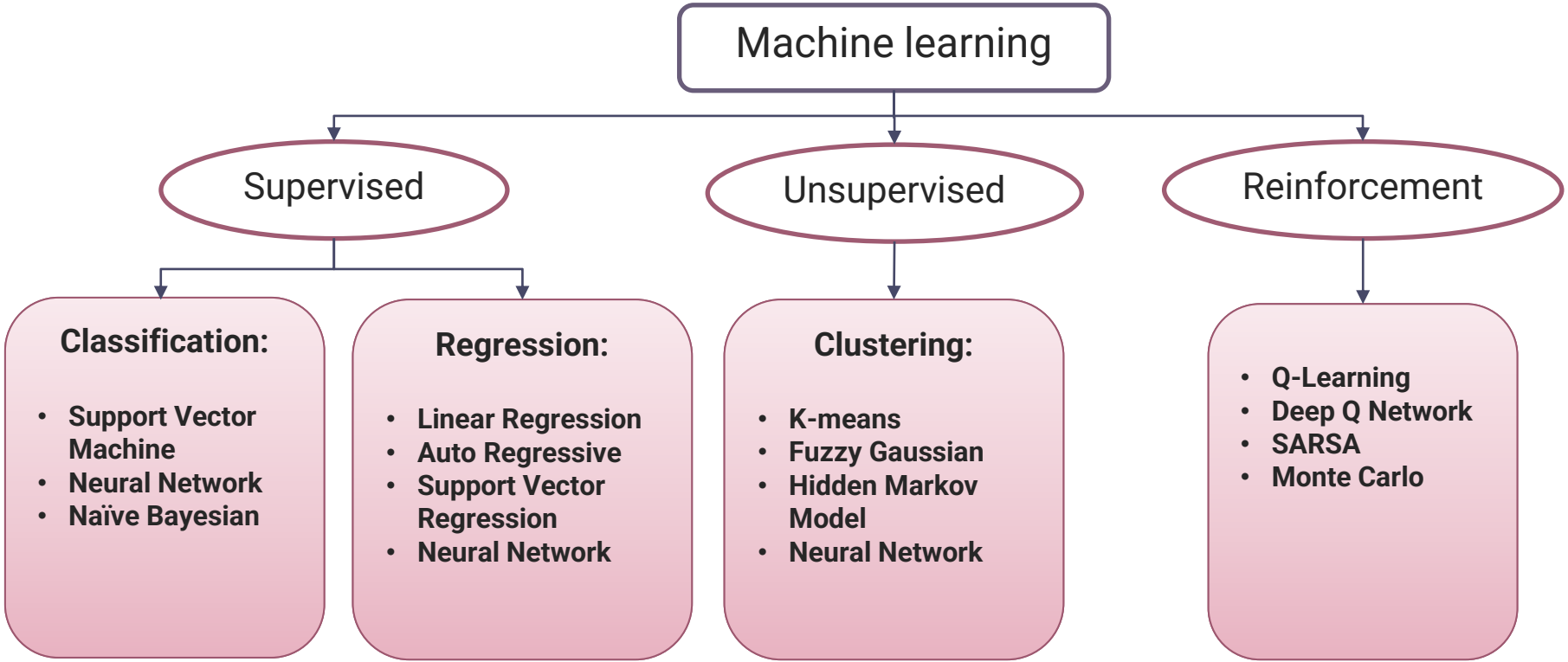
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	No.	Time	Source	Destinatic	Protocol	Length	S port	D port						Duration
2	1	0	192.168.2.5	111.230.24	TCP	74	45623	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
3	10	1.493676	111.230.241.23	192.168.2.	TCP	66	2407	45623	[FIN,	ACK]	Seq=191	Ack=93	Win=7148	1.493676
4	11	1.494509	192.168.2.5	51.38.81.9	TCP	74	48654	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
5	611	7110.126105	192.168.2.5	185.61.14	TCP	74	60987	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
6	612	7110.126177	192.168.2.5	54.39.23.2	TCP	74	51638	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
7	877	10659.59406	192.168.2.5	54.39.23.2	TCP	74	54815	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
8	990	11708.95945	192.168.2.5	54.39.23.2	TCP	66	54815	2407	[FIN,	ACK]	Seq=228	Ack=690	Win=3136	1,049.37
9	992	11709.28096	54.39.23.28	192.168.2.	TCP	66	2407	54815	[FIN,	ACK]	Seq=766	Ack=229	Win=7012	
10	995	12909.05644	192.168.2.5	185.47.12	TCP	74	53935	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
11	998	12913.05784	192.168.2.5	111.230.24	TCP	74	54809	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
12	3197	36593.33376	192.168.2.5	54.39.23.2	TCP	74	41501	2407	[SYN]	Seq=0	Win=2920	Len=0	MSS=1460	
13														
14														
15														
16														
17	Totall time	num of req	complete											
18	129873.3647	8	2											
19														

Data Preprocessing (Total Traffic in Volume)

	A	B	C	D	E	F	G	
1	server	client	app	0-3600	3601-7200	7201-10800	10801-14400	144
2	192.168.2.5	111.230.241.23	53199	0	0	0	0	
3	111.230.241.23	192.168.2.5	53199	0	0	0	0	
4	192.168.2.5	111.230.241.23	22	0	0	0	0	
5	111.230.241.23	192.168.2.5	22	0	0	0	0	
6	192.168.2.5	121.41.16.177	53199	591	0	0	0	
7	121.41.16.177	192.168.2.5	53199	2566	0	0	0	
8	121.41.16.177	192.168.2.5	22	2566	0	0	0	
9	192.168.2.5	121.41.16.177	22	590	0	0	0	
10	121.228.51.251	192.168.2.5	22	2	0	0	0	
11	192.168.2.5	121.228.51.251	22	13	0	0	0	
12	121.42.234.122	192.168.2.5	22	1802	0	0	0	
13	192.168.2.5	121.42.234.122	22	590	0	0	0	
14	121.42.234.122	121.41.16.177	53199	0	0	0	0	
15	121.41.16.177	121.42.234.122	53199	0	0	0	0	
16	121.42.234.122	121.41.16.177	22	0	0	0	0	
17	121.41.16.177	121.42.234.122	22	0	0	0	0	

Sheet1

Machine Learning Algorithms



Implemented ML Models

SVR + PSO

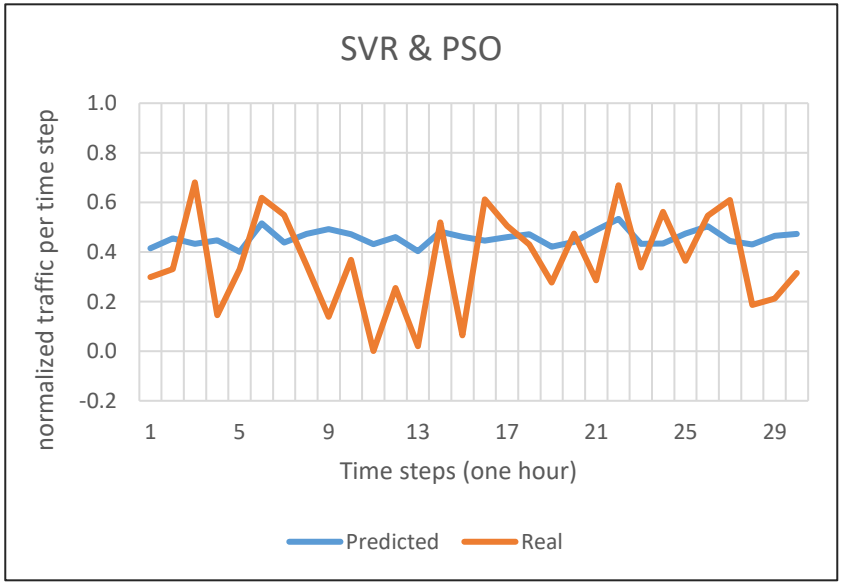
- Support vector regression (SVR) is a type of SVM which can be used for time-series forecasting problems.
- Network traffic prediction is a time series analysis problem, making SVR a suitable solution for this problem.
- A short-term network traffic prediction model based on SVR and particle swarm optimization (PSO) for optimizing SVR parameters(C, γ and ε) to improve network traffic prediction model.

NAR

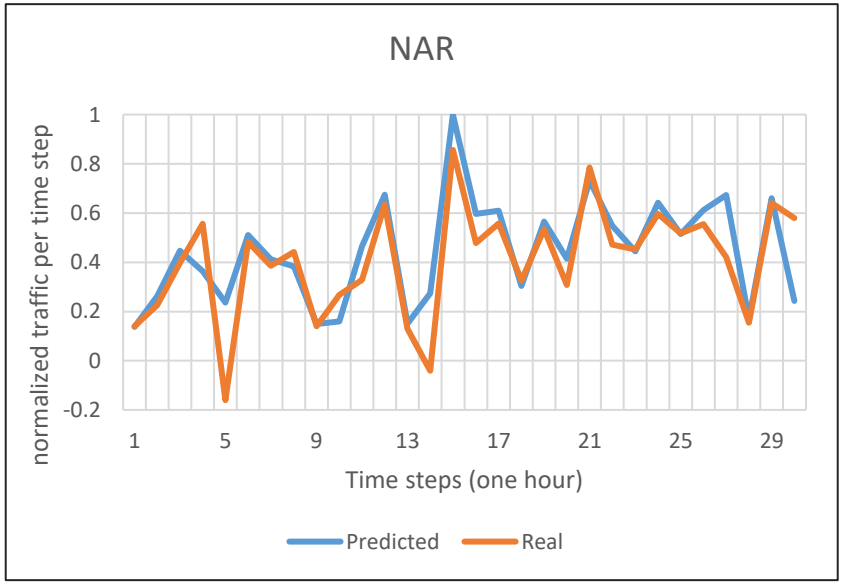
- Nonlinear autoregressive (NAR) neural network
- Predicting a sequence of values in a time series (also known as multistep prediction)
- In NAR prediction, the future values of a time series are predicted only from past values of that series.
- NAR can be written as follows:

$$y(t) = f(y(t-1), \ldots, y(t-d))$$

Results and Evaluation



Number of particles=120 Number of iterations=100
Number of dimensions=3 MSE=0.0408



Number of hidden layers=10 Number of delays=24
Number of iterations=100 MSE=0.0192

Conclusion

Data Preprocess

Converting & labeling the data, deriving a matrix of network features in different types of sessions.

Implementing ML Algorithms

Applying two supervised ML models (optimized **SVR with PSO** & **NAR**) for network traffic prediction.

Results

NAR has shown less MSE in comparison with SVR, making it a better candidate for network traffic prediction.

QUESTIONS

THANK YOU !

Forecasting Model for Service Allocation Network Using Traffic Recognition

Ghazal Rahmanian

Mohammad Yekta

HadiShahriar Shahhoseini

AmirHossein Jafari Pozveh

Iran University of Science and Technology