Data Mining and Machine Learning for Analysis of Network Traffic

Ljiljana Trajković
ljilja@cs.sfu.ca

Communication Networks Laboratory
http://www.sfu.ca/~ljilja/cnl
School of Engineering Science
Simon Fraser University, Vancouver
British Columbia, Canada

Simon Fraser University
Burnaby Campus

1965
SFU 2015

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

Measurements of Network Traffic

- Traffic measurements:
  - help understand characteristics of network traffic
  - are basis for developing traffic models
  - are used to evaluate performance of protocols and applications
- Traffic analysis:
  - provides information about the network usage
  - helps understand the behavior of network users
- Traffic prediction:
  - important to assess future network capacity requirements
  - used to plan future network developments

Data Collections

- Data collected from networks are used to:
  - evaluate network performance
  - characterize and model traffic
  - identify trends in the evolution of the Internet topology
  - classify traffic and network anomalies

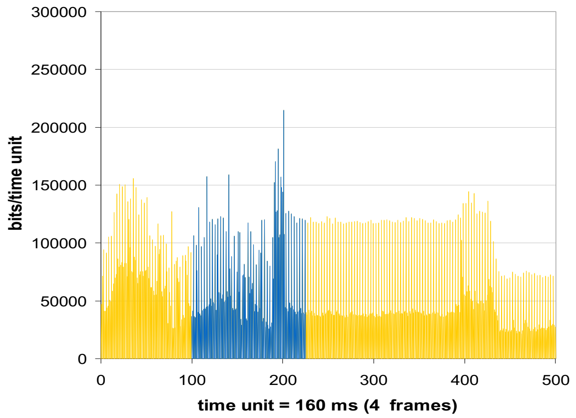
Traffic Modeling: Self-Similarity

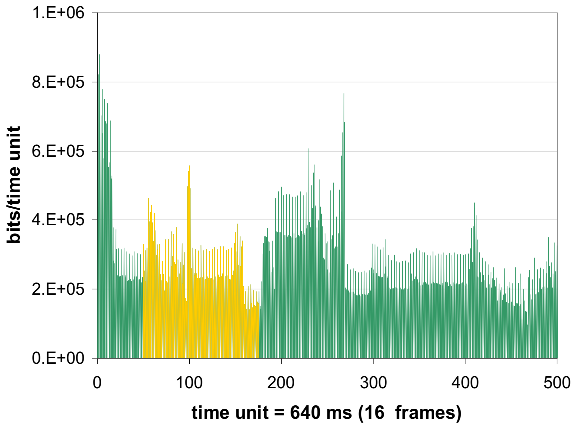
- Self-similarity implies a ‘‘fractal-like’’ behavior:
  data on various time scales have similar patterns
- A wide-sense stationary process $X(n)$ is called (exactly second order) self-similar if its autocorrelation function satisfies:
  - $r^{(m)}(k) = r(k)$, $k \geq 0$, $m = 1, 2, \ldots, n$,

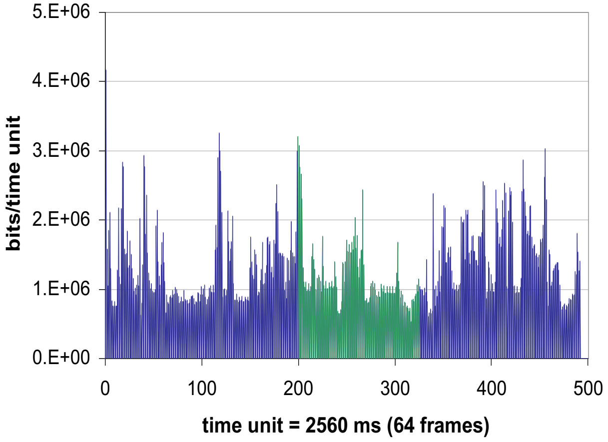
    where m is the level of aggregation

Self-Similarity: Influence of Time-Scales

- Genuine MPEG traffic trace



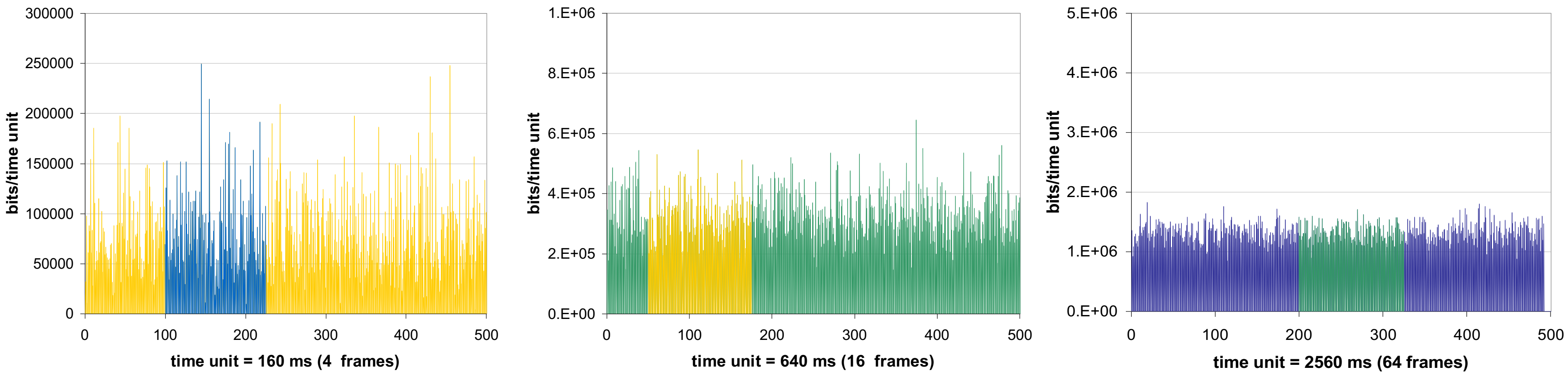




W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, “On the self-similar nature of Ethernet traffic (extended version),” IEEE/ACM Trans. Netw., vol. 2, no 1, pp. 1-15, Feb. 1994.

Self-Similarity: Influence of Time-Scales

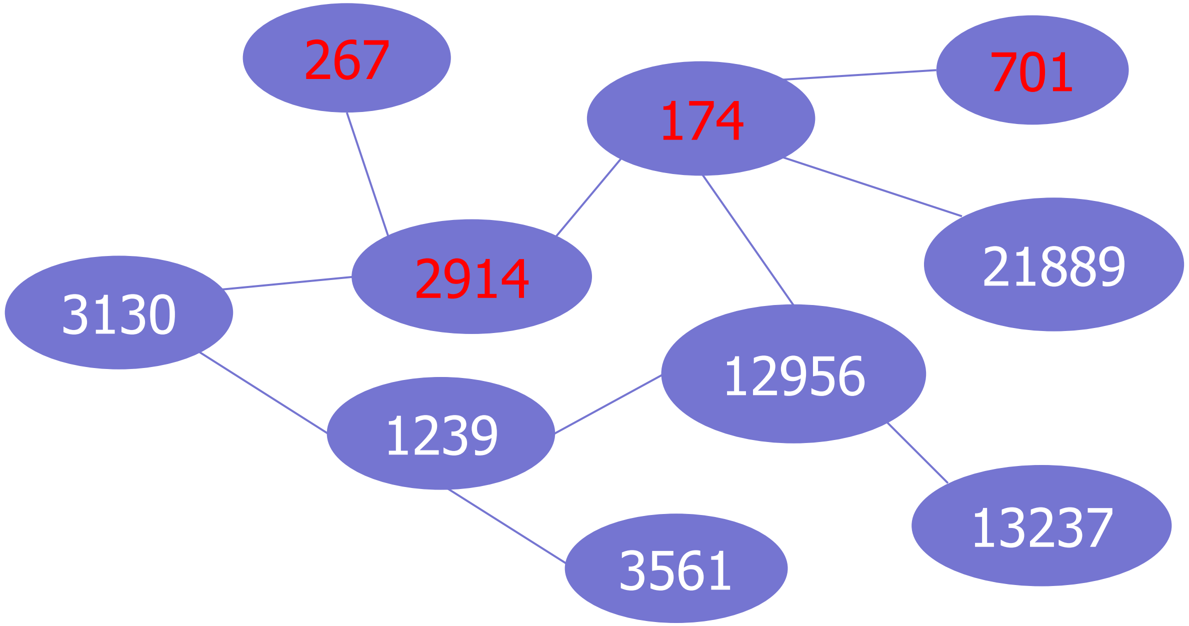
- Synthetically generated Poisson model



W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, “On the self-similar nature of Ethernet traffic (extended version),” IEEE/ACM Trans. Netw., vol. 2, no 1, pp. 1-15, Feb. 1994.

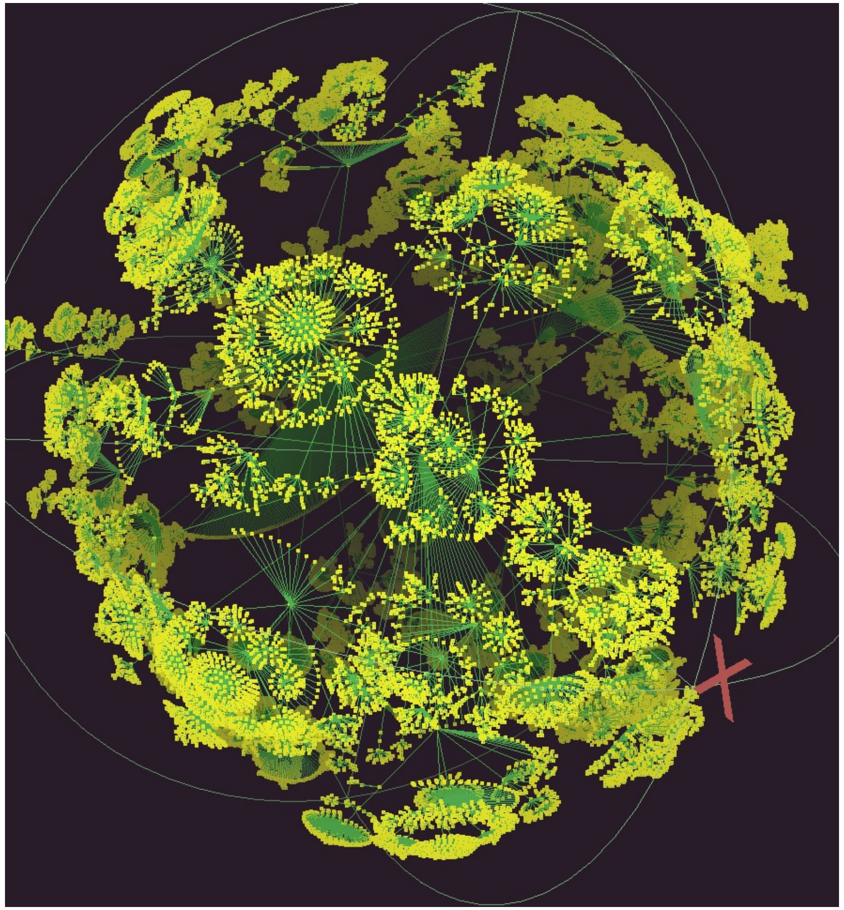
Internet Topology at AS Level

- Collected data from Border Gateway Protocols (BGP) routing tables are used to infer the Internet topology



G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, “Power-laws and the AS-level Internet topology,” IEEE/ACM Trans. Networking, vol. 11, no. 4, pp. 514–524, Aug. 2003.

The Internet Topology: Scale Free Graphs



http://www.caida.org/home/

Ihr: 535,102 nodes and 601,678 links

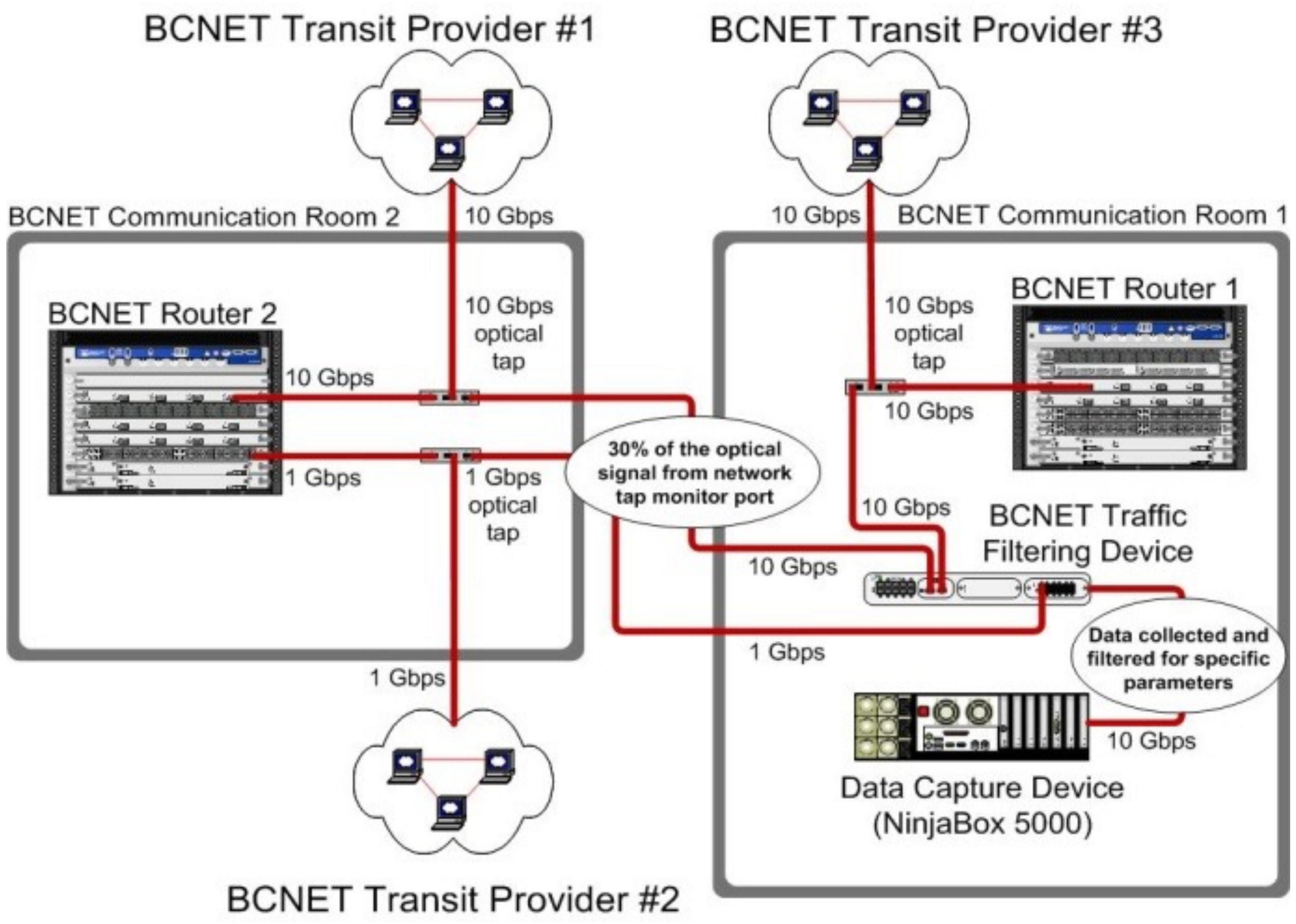
Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, the Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

Case Study: BCNET

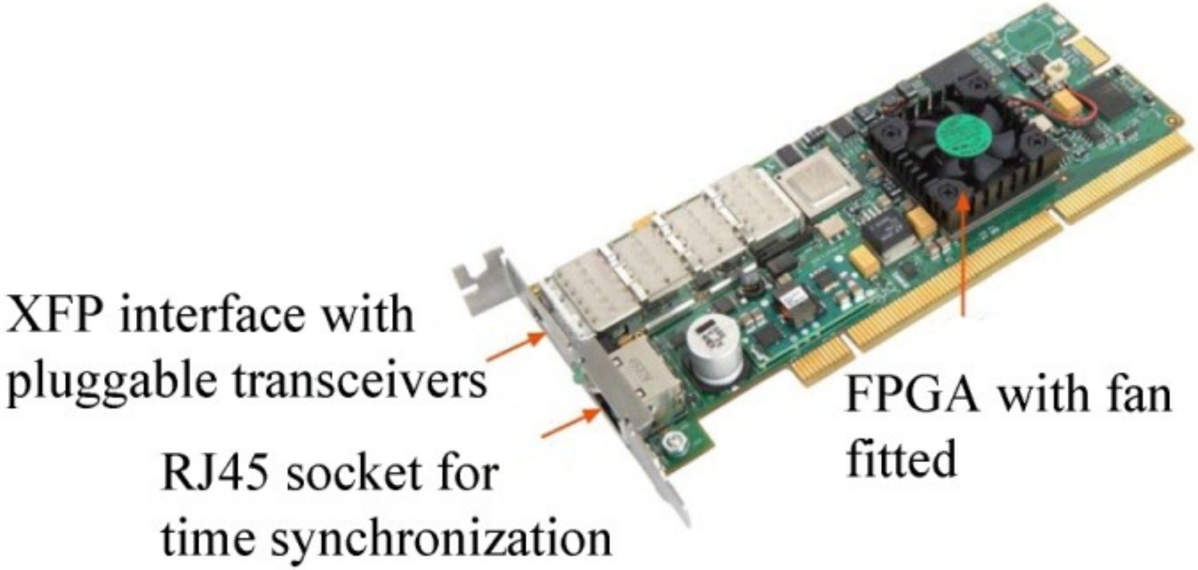
- BCNET is the hub of advanced telecommunication network in British Columbia, Canada that offers services to research and higher education institutions
- The BCNET network is high-speed fiber optic research network
- British Columbia's network extends to 1,400 km and connects Kamloops, Kelowna, Prince George, Vancouver, and Victoria

BCNET Packet Capture



Network Monitoring and Analyzing

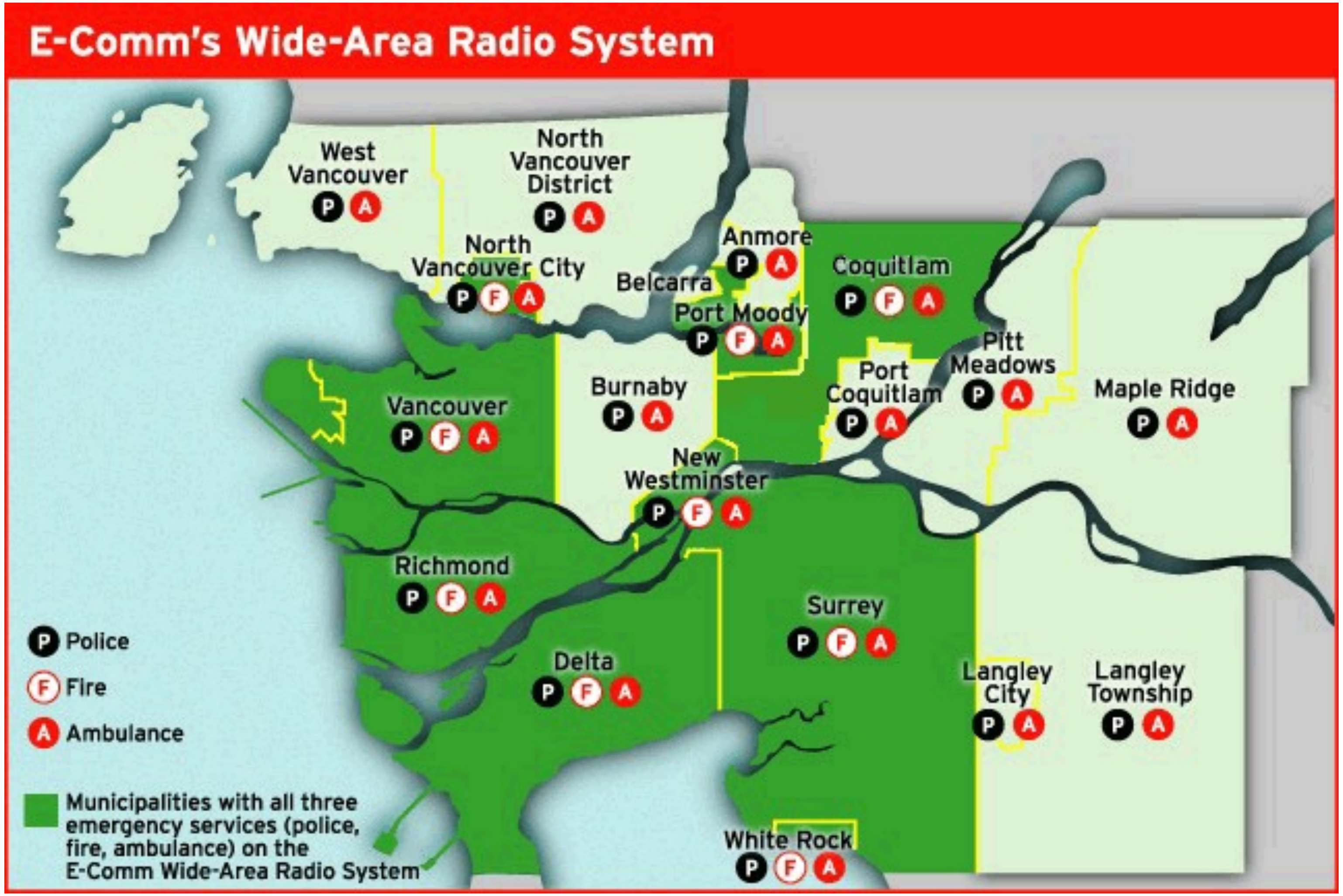
- Endace Data Acquisition and Generation (DAG) 5.2X card
- Captures and transmits traffic and has time-stamping capability
- DAG 5.2X is a single port Peripheral Component Interconnect Extended (PCIx) card and is capable of capturing on average Ethernet traffic of 6.9 Gbps



Case Study: E-Comm Network

- E-Comm network: an operational trunked radio system serving as a regional emergency communication system
- The E-Comm network enables both voice and data transmissions
- Voice traffic accounts for over 99% of network traffic
- More than 85% of calls are group calls
- A distributed event log database records every event occurring in the network:
  - call establishment
  - channel assignment
  - call drop
  - emergency call

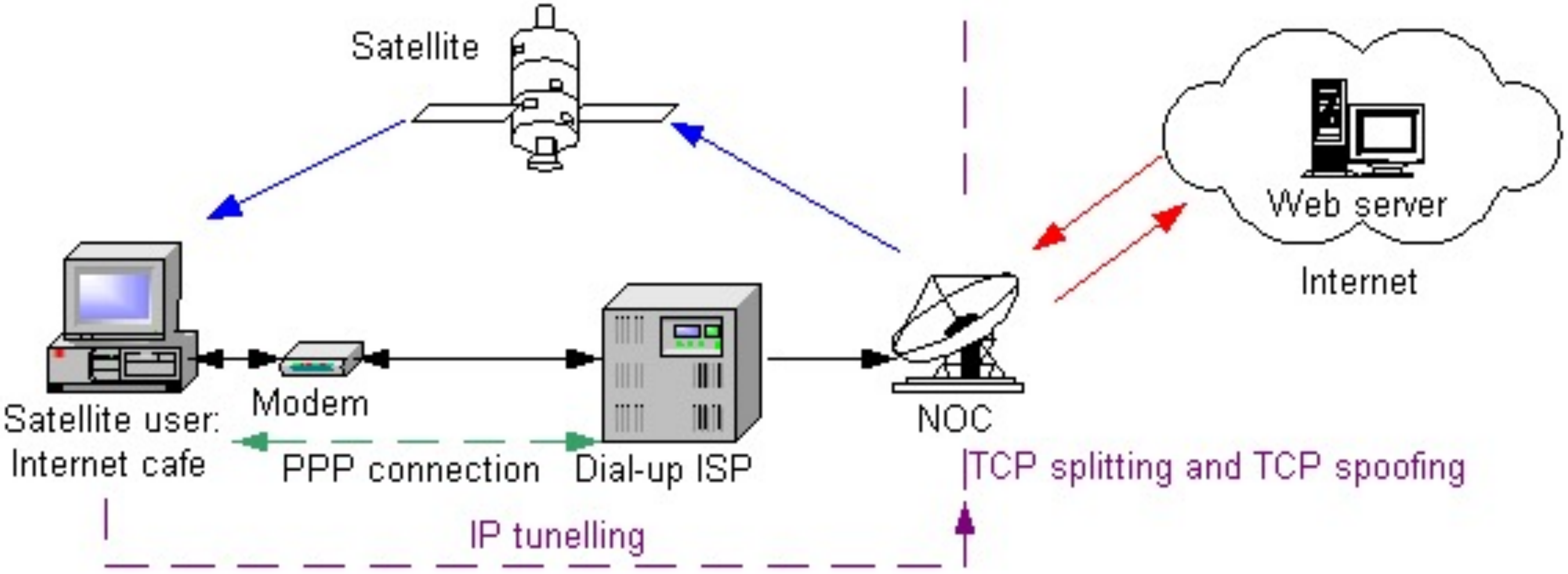
E-Comm Network



Case Study: ChinaSat DirecPC System

- ChinaSat hybrid satellite network
  - Employs geosynchrous satellites deployed by Hughes Network Systems Inc.
  - Provides data and television services:
    - DirecPC (Classic): unidirectional satellite data service
    - DirecTV: satellite television service
    - DirecWay (Hughnet): bi-directional satellite data service that replaces DirecPC
  - DirecPC transmission rates:
    - 400 kb/s from satellite to user
    - 33.6 kb/s from user to network operations center (NOC) using dial-up
  - Improves performance using TCP splitting with spoofing

ChinaSat DirecPC System



Traffic Anomalies and Intrusions

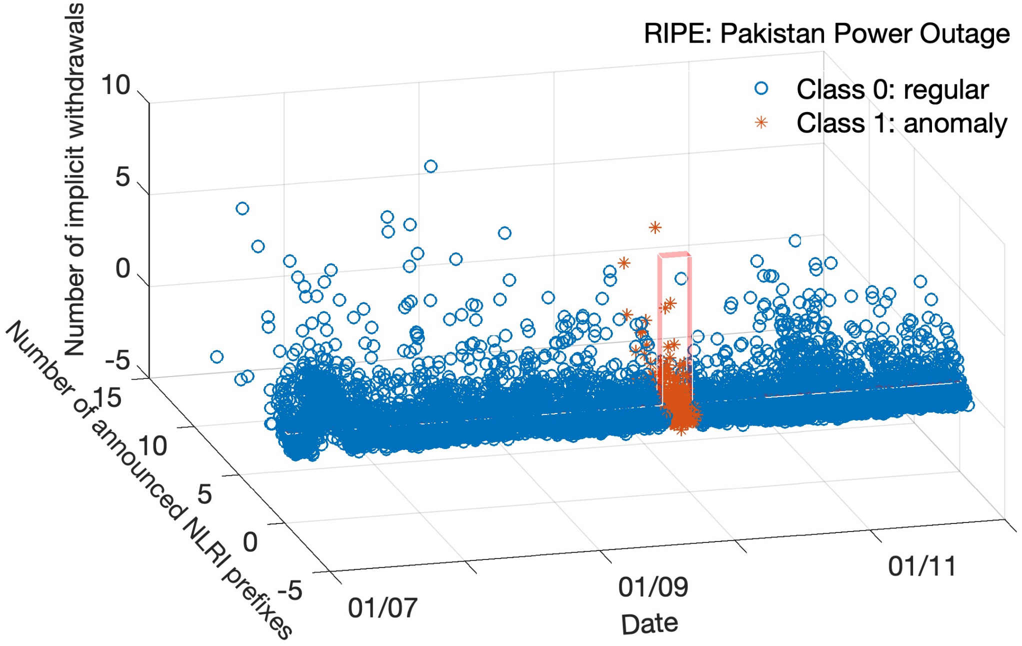
- Anomalies affect performance of the Internet Border Gateway Protocol
  - Computer worms and viruses:
    - Slammer (2003), Nimda (2001), Code Red (2001)
  - Electrical failures:
    - Moscow blackout (2005) and Pakistan power outage (2021)
  - Ransomware attacks:
    - WannaCrypt (2017) and WestRock (2021)
  - Internet Protocol (IP) prefix hijacks, miss-configurations

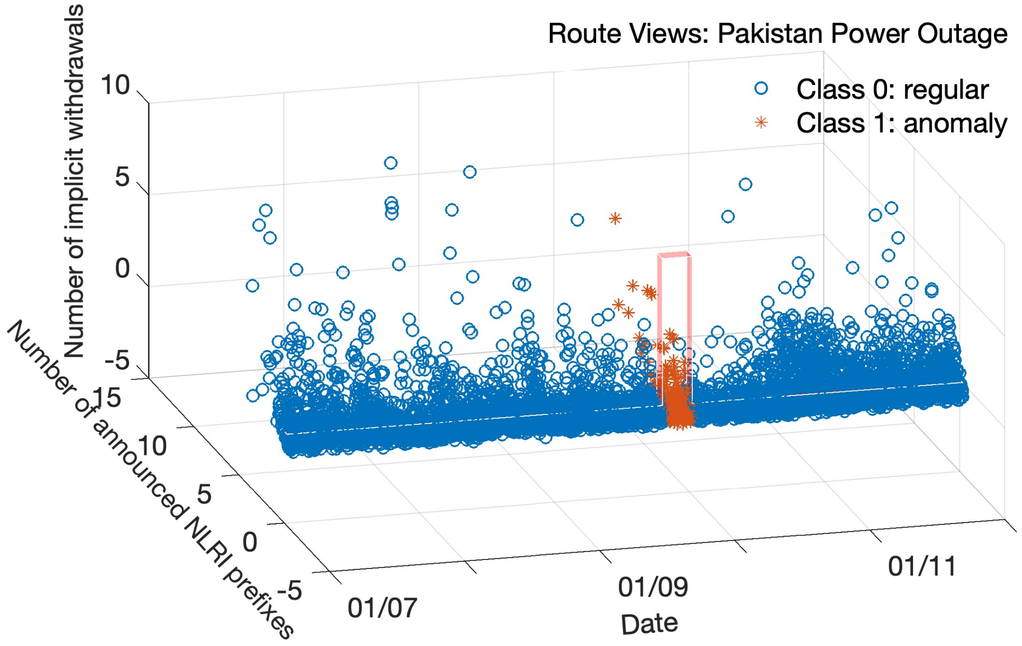
Network Traffic Datasets

- Internet Border Gateway Protocol (BGP) anomalies:
  - Computer worms and viruses:
    - Code Red (2001), Nimda (2001), Slammer (2003)
  - Electrical failures:
    - Moscow blackout (2005) and Pakistan power outage (2021)
  - Ransomware attacks:
    - WannaCrypt (2017) and WestRock (2021)
  - Internet Protocol (IP) prefix hijacks, miss-configurations
- Collection sites:
  - Réseaux IP Européens (RIPE)
  - Route Views

BGP Dataset: Pakistan Power Outage (2021)

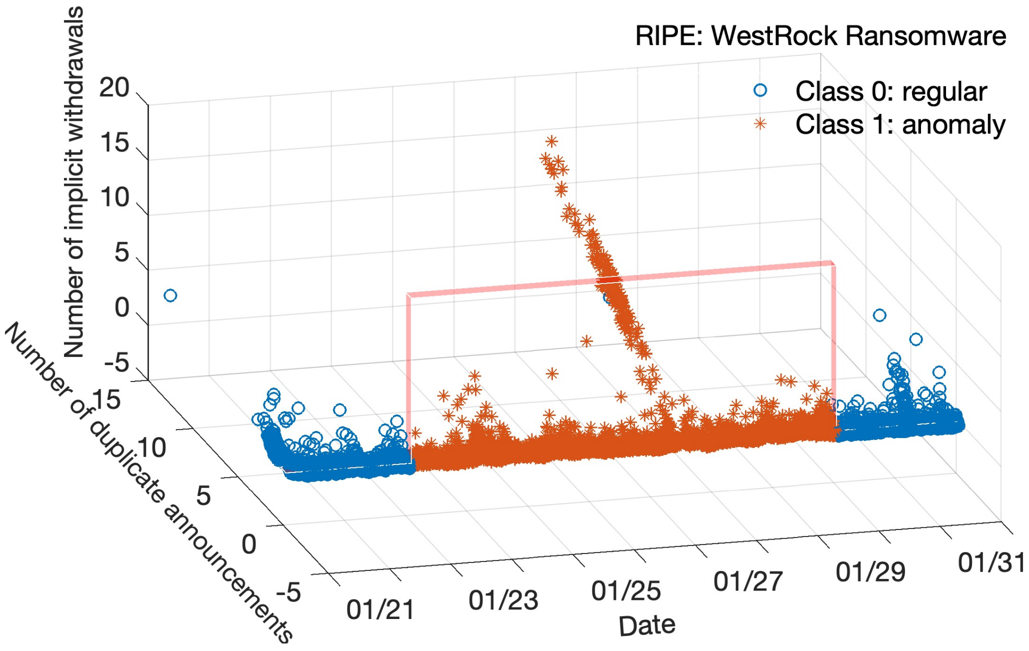
- Number of announced NLRI prefixes vs. number of implicit withdrawals vs. date:

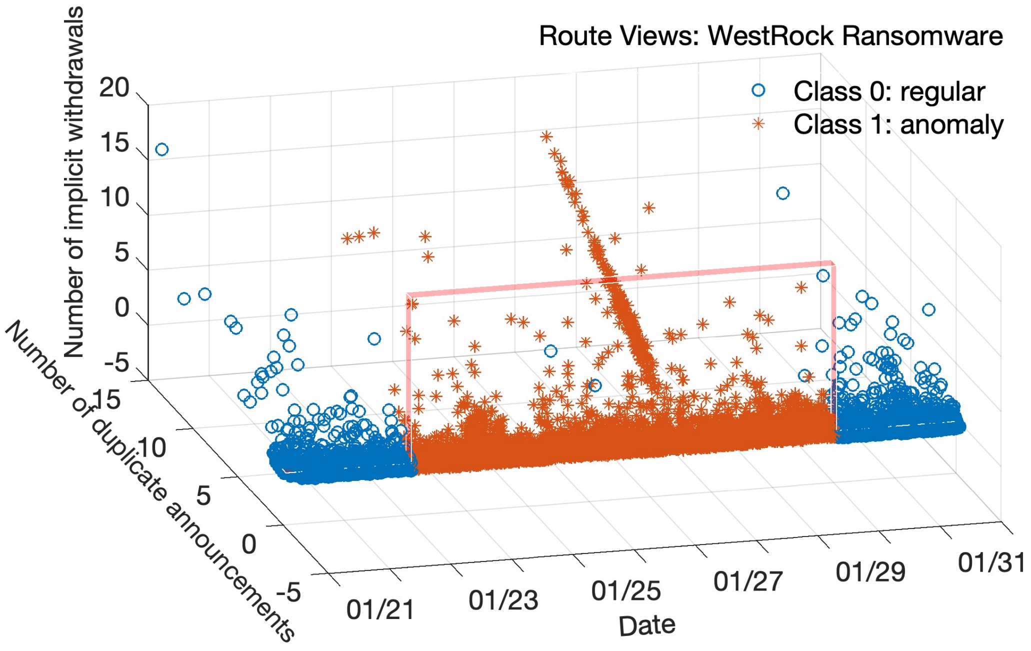




BGP Dataset: WestRock Ransomware Attack (2021)

- Number of announced NLRI prefixes vs. number of implicit withdrawals vs. date:





BGP Datasets: Internet worms

- Slammer, Nimda, Code Red:

Collection site	Dataset	Regular (min)	Anomaly (min)	Regular (training)	Anomaly (training)	Regular (test)	Anomaly (test)	Start	End
RIPE	Code Red	6,600	600	3,679	361	2,921	239	17.07.2001 00:00:00	21.07.2001 23:59:59
	Nimda	7,308	1,301	3,673	827	3,635	474	16.09.2001 00:00:00	21.09.2001 23:59:59
	Slammer	6,331	869	3,210	530	3,121	339	23.01.2003 00:00:00	27.01.2003 23:59:59
Route Views	Slammer	6,319	869	3,198	530	3,121	339	23.01.2003 00:00:00	27.01.2003 23:59:59

Route Views data collection began in 2003.

BGP Datasets: Power Blackouts and Outages

Collection site	Dataset	Regular (min)	Anomaly (min)	Regular (training)	Anomaly (training)	Regular (test)	Anomaly (test)	Start	End
RIPE	Moscow blackout	6,960	240	3,120	180	3,840	60	23.05.2005 00:00:00	27.05.2005 23:59:59
	Pakistan power outage	6,880	320	4,000	200	2,880	120	07.01.2021 00:00:00	11.01.2021 23:59:59
Route Views	Moscow blackout	6,865	130	3,075	85	3,790	45	23.05.2005 00:00:00	27.05.2005 23:59:59
	Pakistan power outage	6,880	320	4,000	200	2,880	120	07.01.2021 00:00:00	11.01.2021 23:59:59

BGP Datasets: Ransomware Attacks

Collection site	Dataset	Regular (min)	Anomaly (min)	Regular (training)	Anomaly (training)	Regular (test)	Anomaly (test)	Start	End
RIPE/ Route Views	WannaCrypt	5,760	5,760	2,880	3,420	2,880	2,340	10.05.2017 00:00:00	17.05.2017 23:59:59
	WestRock ransomware	5,832	10,008	2,952	6,008	2,880	4,000	21.01.2021 00:00:00	31.01.2021 23:59:59

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models:
  - Deep learning: multi-layer recurrent neural networks
  - Broad learning system
  - Gradient boosting decision trees
- Experimental procedure
- Performance evaluation
- Conclusions and references

Machine Learning Algorithms

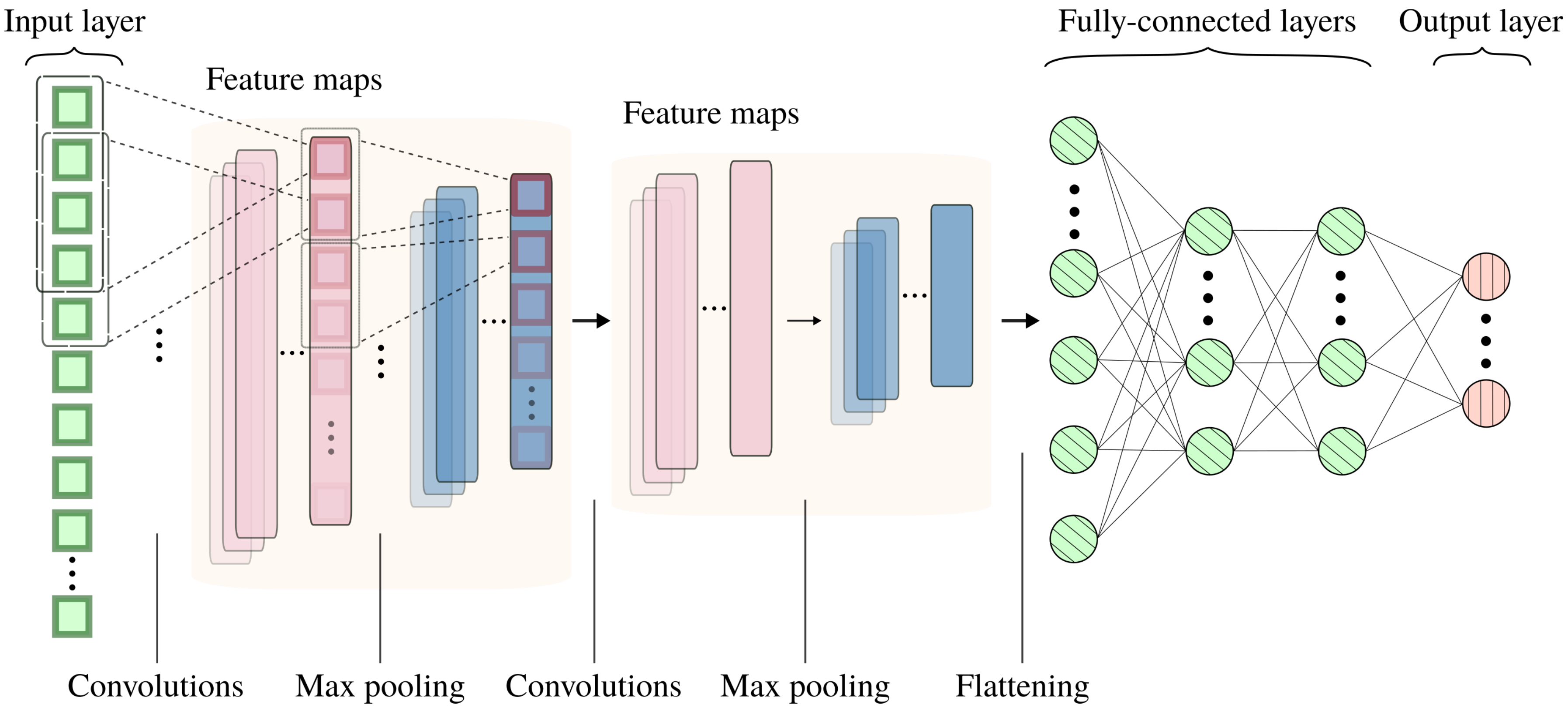
- Network intrusion detection systems employ diverse:
  - Deep learning algorithms:
    - Convolutional neural networks: CNNs
    - Recurrent neural networks: RNNs
    - Deep belief networks
    - Autoencoders
  - Boosting algorithms:
    - AdaBoost
    - Gradient boosting decision trees

Machine Learning Algorithms

- Supervised machine learning algorithms:
  - Support vector machine: SVM
  - Long short-term memory: LSTM and Bi-LSTM
  - Gated recurrent unit: GRU and Bi-GRU
  - Gradient Boosting Decision Trees (GBDT):
    - XGBoost
    - LightGBM
    - CatBoost
  - Broad learning system: BLS and its extensions

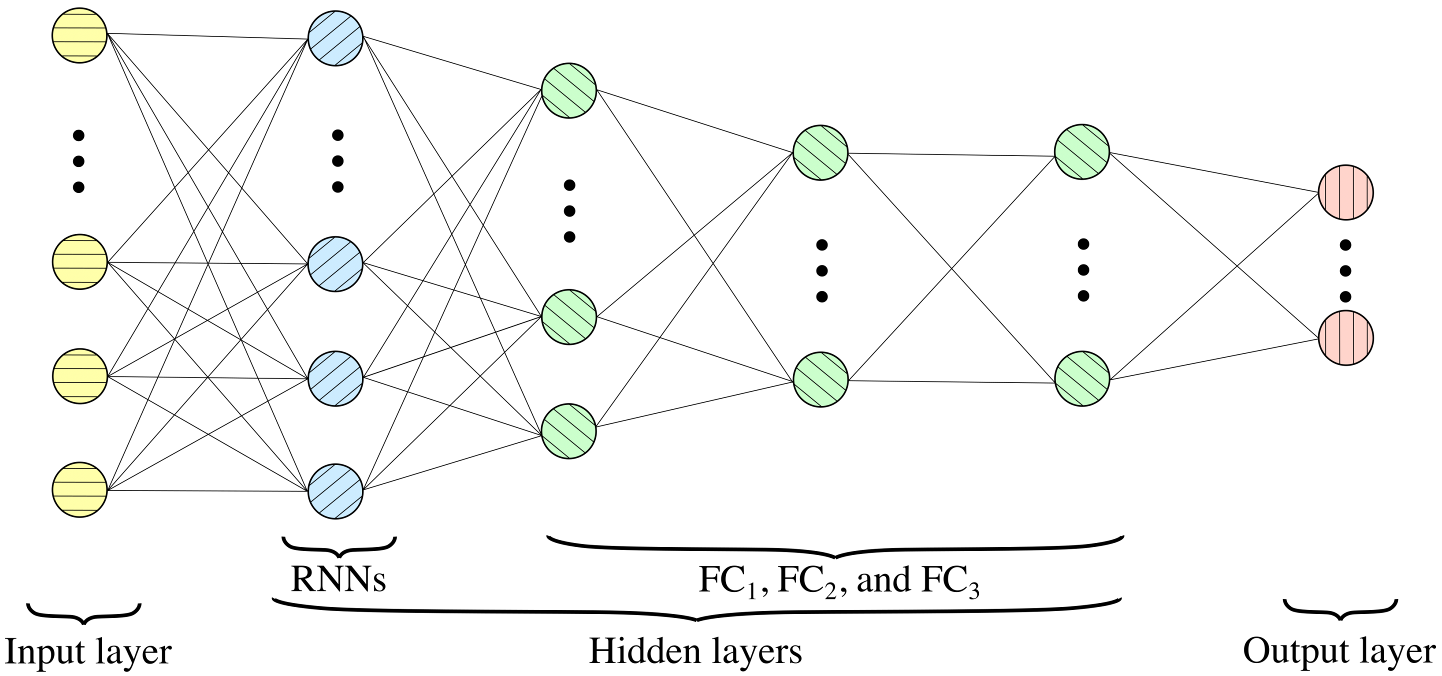
Convolutional Neural Network

- The high-level structure of a CNN using 1-dimensional input data:



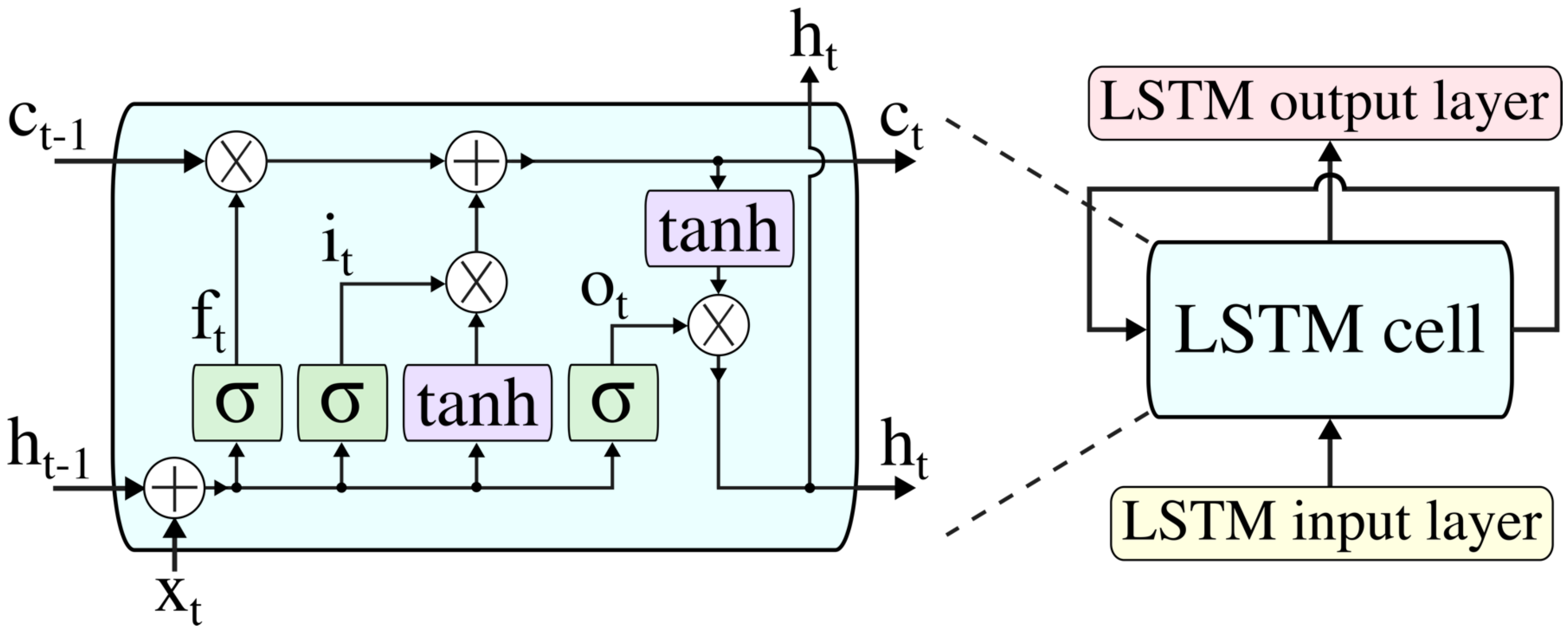
Deep Learning Neural Network

- 37 (BGP)/109 (NSL-KDD) RNNs, 80 FC_1, 32 FC_2, and 16 FC_3 fully connected (FC) hidden nodes:



Long Short-Term Memory

- Repeating module for the Long Short-Term Memory (LSTM) neural network:



Long Short-Term Memory: LSTM

- The outputs of the forget gate f_t, the input gate i_t, and the output gate o_t at time t are:

$$f_t = \sigma\left(W_{if}x_t + b_{if} + U_{hf}h_{t-1} + b_{hf}\right)$$
$$i_t = \sigma(W_{ii}x_t + b_{ii} + U_{hi}h_{t-1} + b_{hi})$$
$$o_t = \sigma(W_{io}x_t + b_{io} + U_{ho}h_{t-1} + b_{ho}),$$

where:

$\sigma(\cdot)$: logistic sigmoid function

x_t: current input vector

h_{t-1}: previous output vector

W_{if}, U_{hf}, W_{ii}, U_{hi}, W_{io}, and U_{ho}: weight matrices

b_{if}, b_{hf}, b_{ii}, b_{hi}, b_{io}, and b_{ho}: bias vectors

Long Short-Term Memory: LSTM

- Output i_t of the input gate decides if the information will be stored in the cell state. The sigmoid function is used to update the information.
- Cell state c_t:

$$c_t = f_t * c_{t-1} + i_t * tanh(W_{ic}x_t + b_{ic} + U_{hc}h_{t-1} + b_{hc}),$$

where:

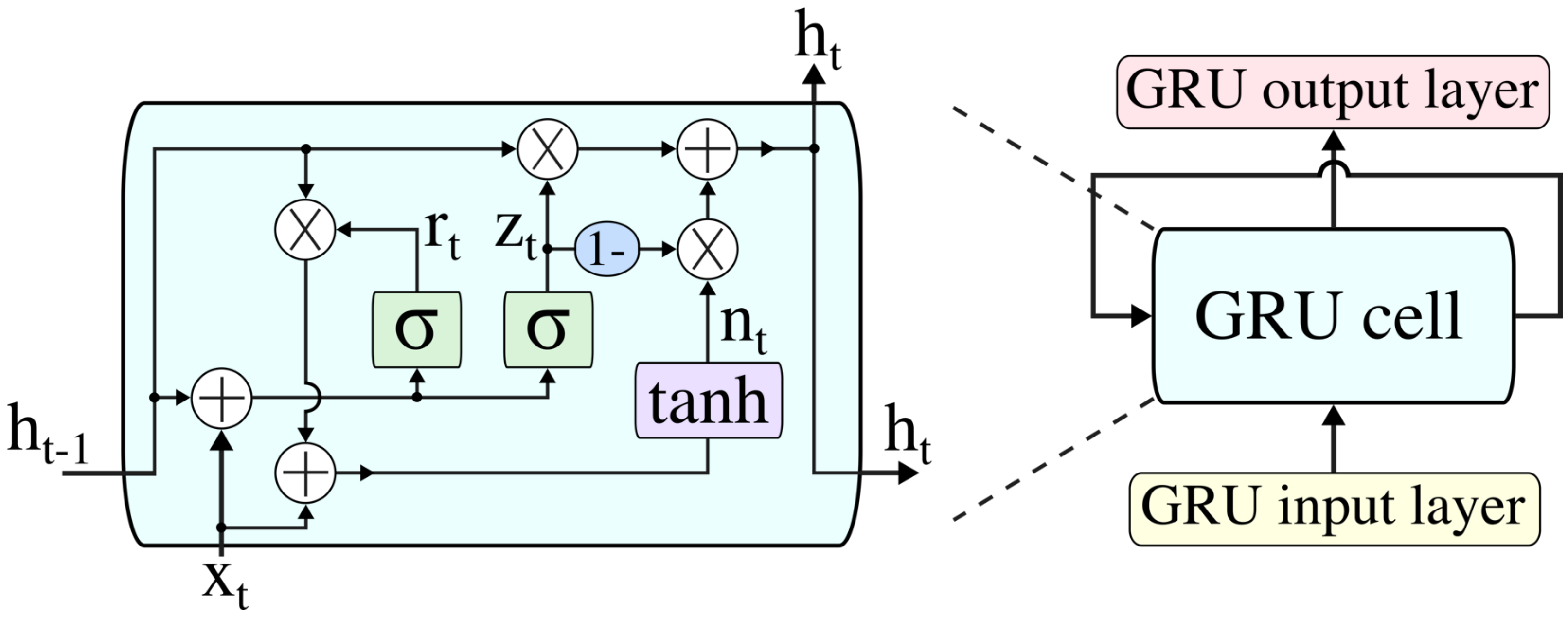
  - $*$ denotes element-wise multiplications
  - $tanh$ function: used to create a vector for the next cell state

- Output of the LSTM cell:

$$h_t = o_t * tanh(c_t)$$

Gated Recurrent Unit

- Repeating module for the Gated Recurrent Unit (GRU) neural network:



Gated Recurrent Unit: GRU

- The outputs of the reset gate r_t and the update gate z_t at time t:

$$r_t = \sigma(W_{ir}x_t + b_{ir} + U_{hr}h_{t-1} + b_{hr})$$
$$z_t = \sigma(W_{iz}x_t + b_{iz} + U_{hz}h_{t-1} + b_{hz}),$$

where:

  - σ: sigmoid function
  - x_t: input
  - h_{t-1} is the previous output of the GRU cell
  - W_{ir}, U_{hr}, W_{iz}, and U_{hz}: weight matrices
  - b_{ir}, b_{hr}, b_{iz} +, and b_{hz} : bias vectors

Gated Recurrent Unit: GRU

- Output of the GRU cell:

$$h_t = (1 - z_t) * n_t + z_t * h_{t-1},$$

where n_t:

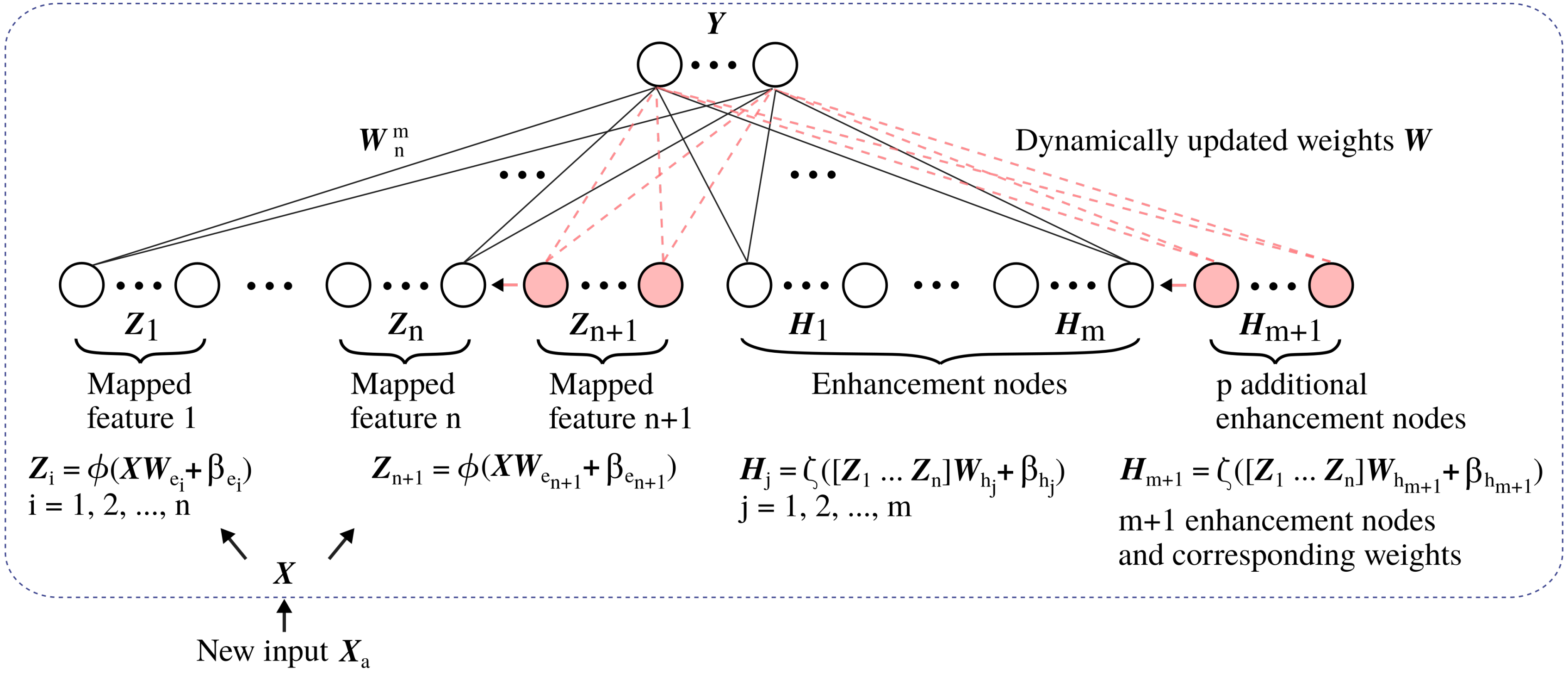
  - $n_t = tanh(W_{in}x_t + b_{in} + r_t * (U_{hn}h_{t-1} + b_{hn}))$
  - W_{in} and U_{hn}: weight matrices
  - b_{in} and b_{hn}: bias vectors

Roadmap

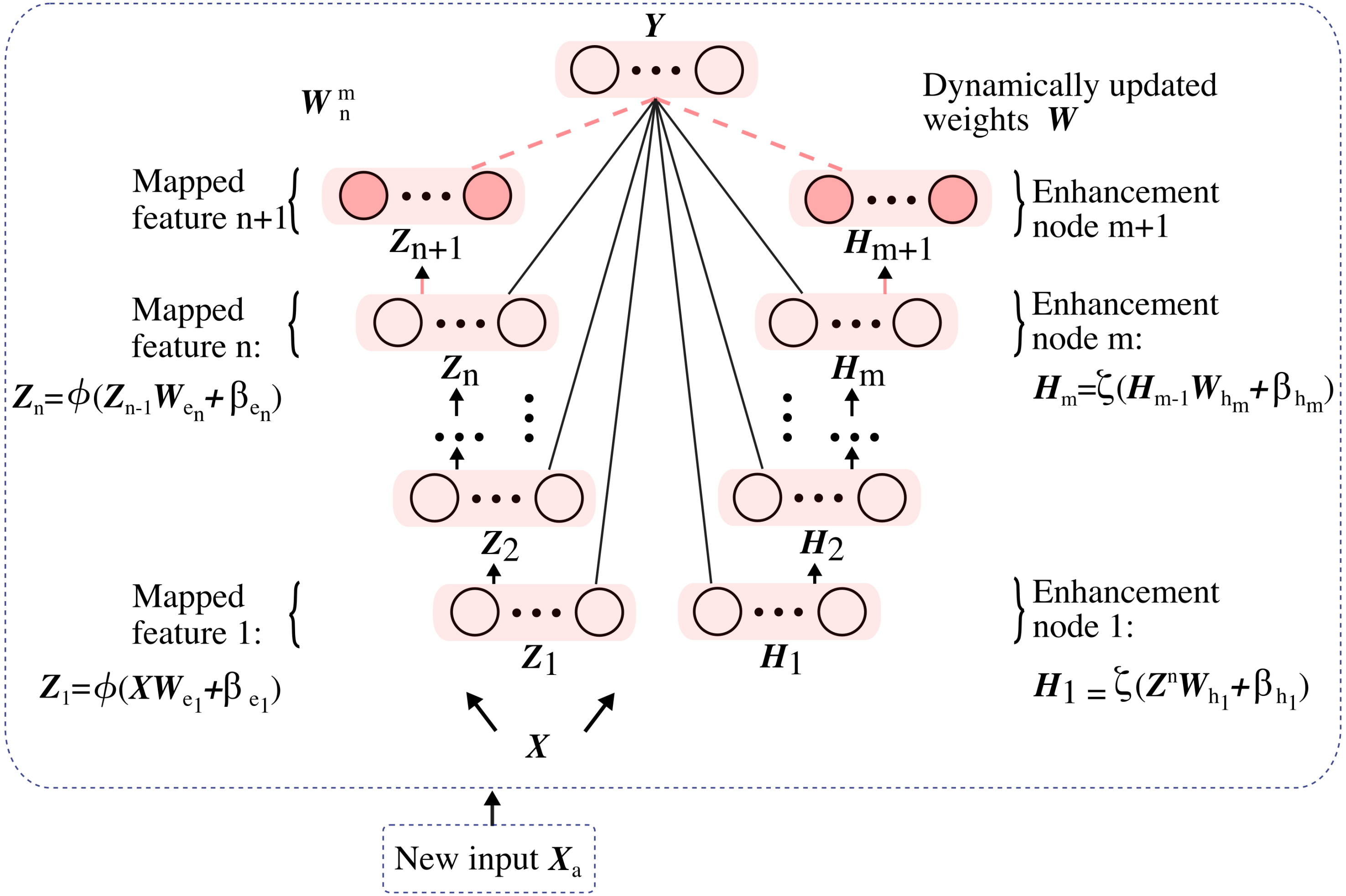
- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models:
  - Deep learning: multi-layer recurrent neural networks
  - Broad learning system
  - Gradient boosting decision trees
- Experimental procedure
- Performance evaluation
- Conclusions and references

Broad Learning System

- Broad Learning System (BLS) algorithm with increments of mapped features, enhancement nodes, and new input data:

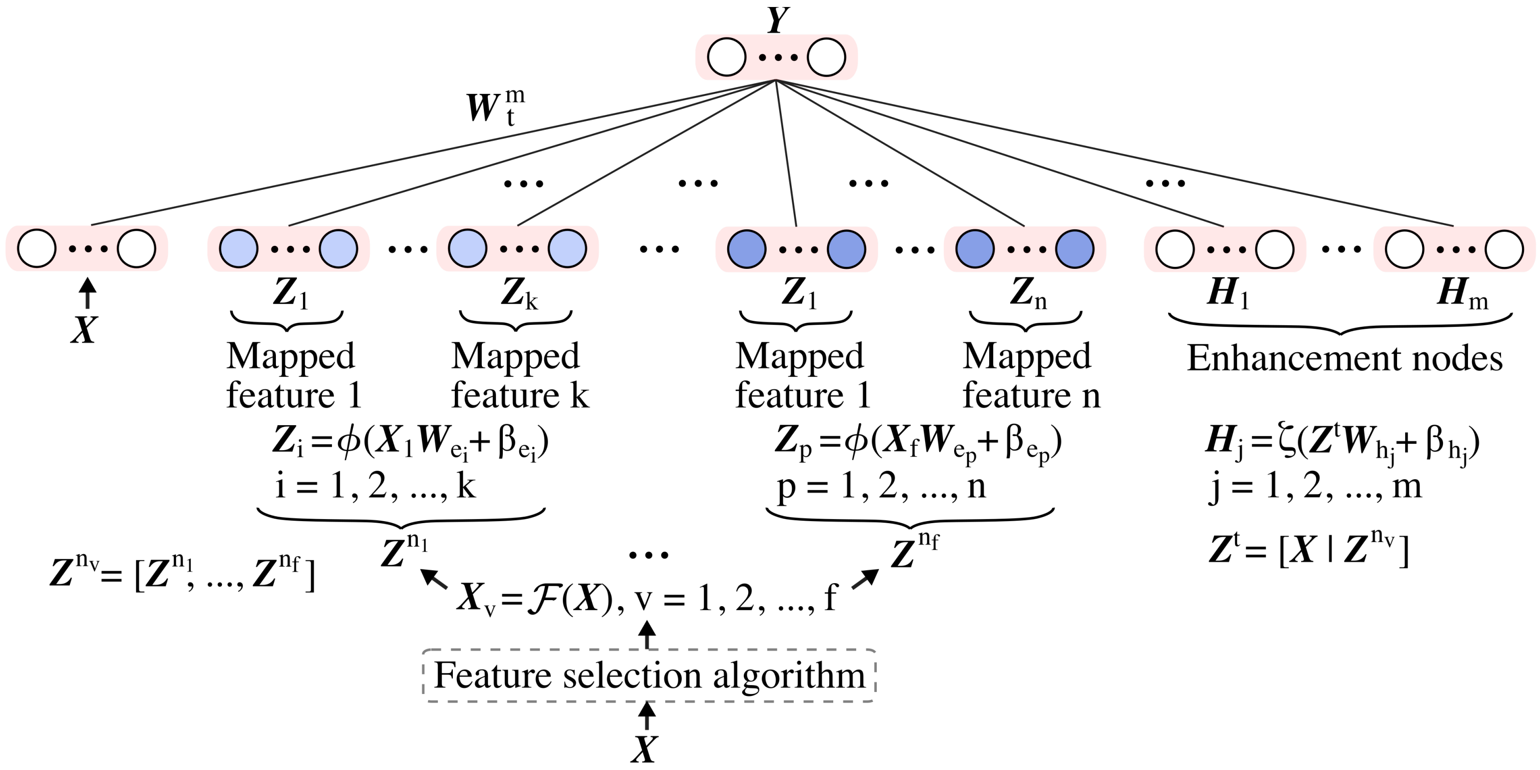


Cascades with Incremental Learning



Variable Features Broad Learning System

- VFBLS



Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models:
  - Deep learning: multi-layer recurrent neural networks
  - Broad learning system
  - Gradient boosting decision trees
- Experimental procedure
- Performance evaluation
- Conclusions and references

Gradient Boosting Decision Trees

GBDT algorithms:

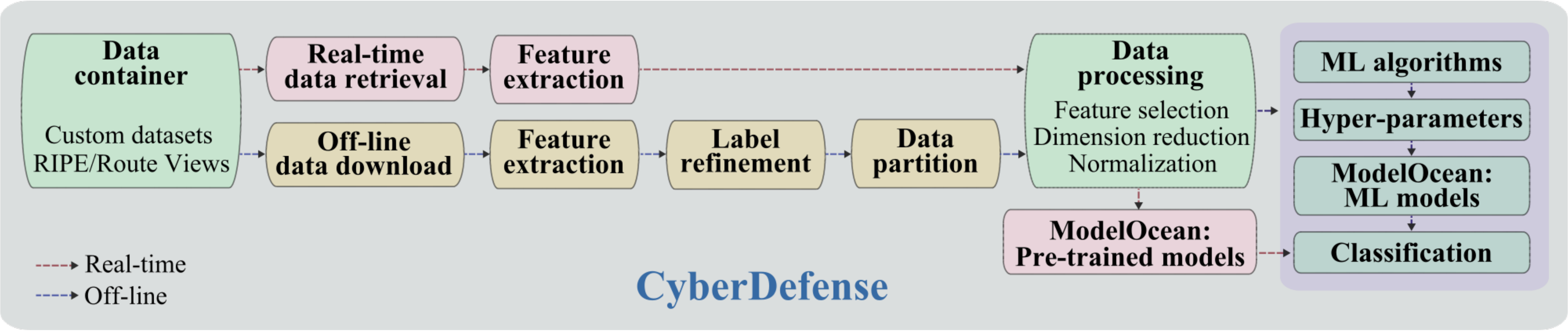
- XGBoost: eXtreme gradient boosting
- LightGBM: light gradient boosting machine
- CatBoost: categorical boosting

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

CyberDefense

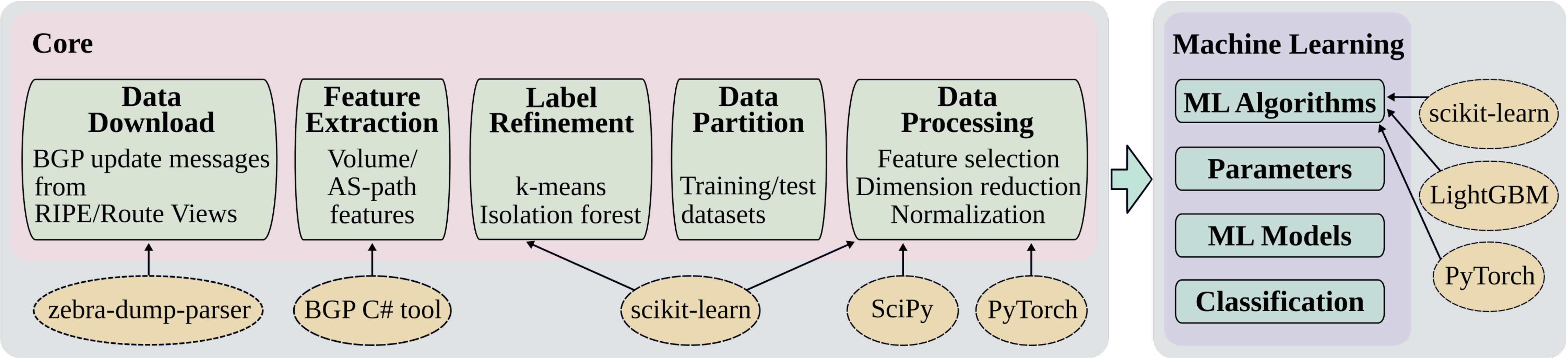
- Architecture:



CyberDefense: Tool for Detecting Network Anomalies and Intrusions, https://github.com/zhida-li/cyberDefense

BGPGuard

- Architecture:



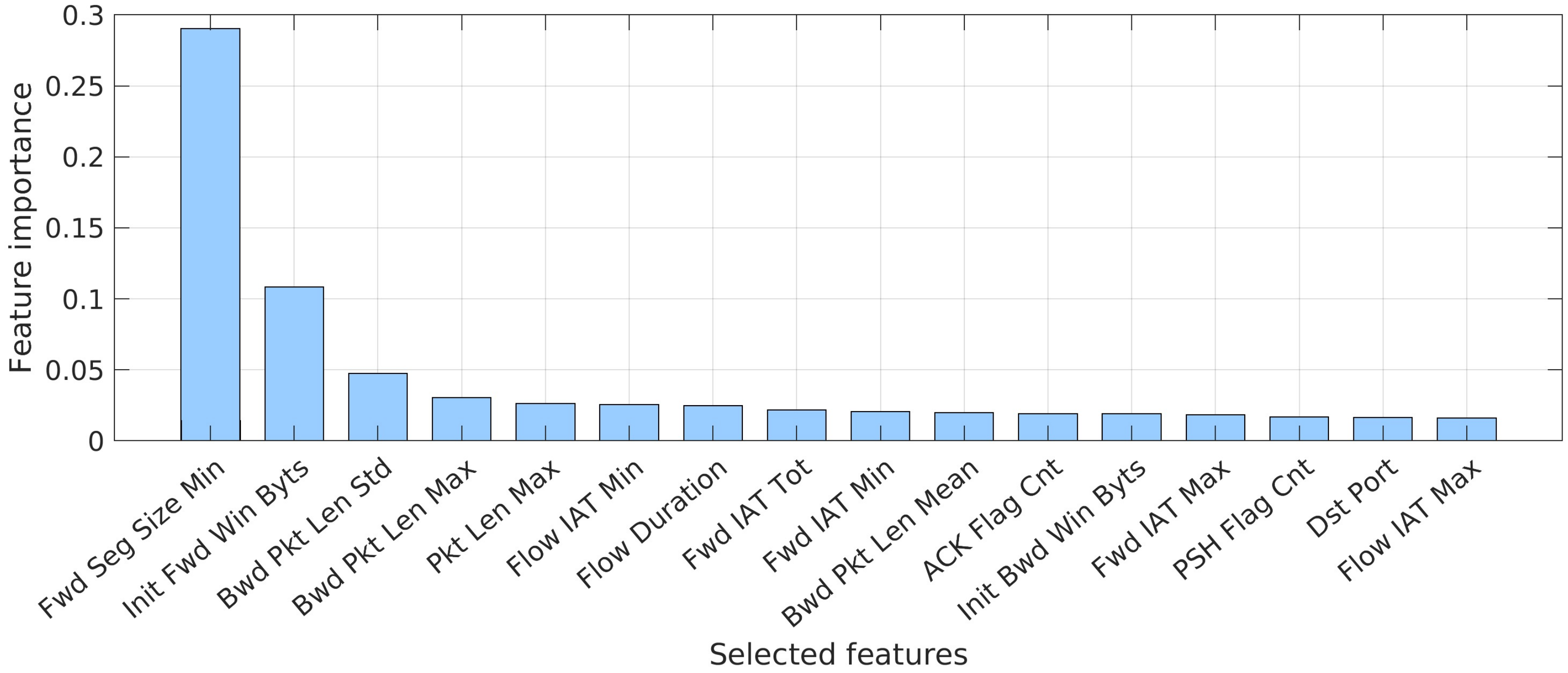
BGPGuard: a BGP Anomaly Detection Tool,
https://github.com/zhida-li/BGPGuard

Experimental Procedure

- Step 1: Normalize training and test datasets.
- Step 2: Train the models using 10-fold validation and tune model parameters.
- Step 3: Test the best models.
- Step 4: Evaluate models based on:
  - accuracy
  - F-score
  - precision
  - sensitivity
  - confusion matrix
  - training time

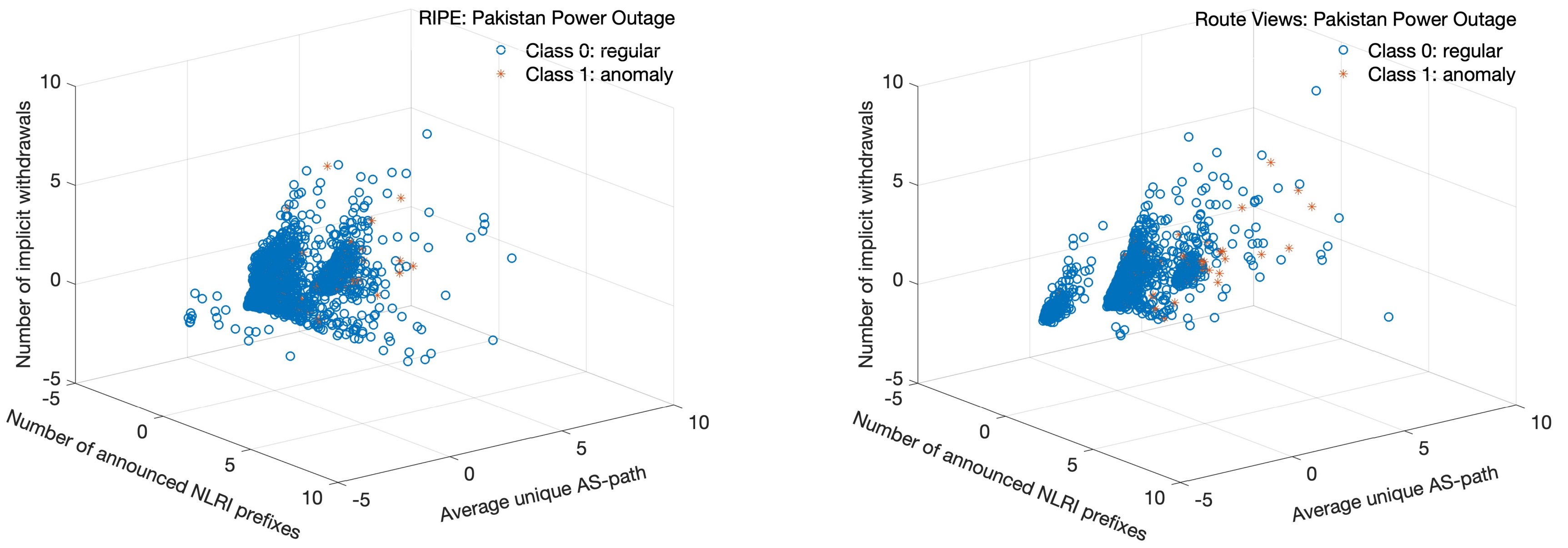
Most Relevant Features

- CSE-CIC-IDS2018: 16 most relevant features



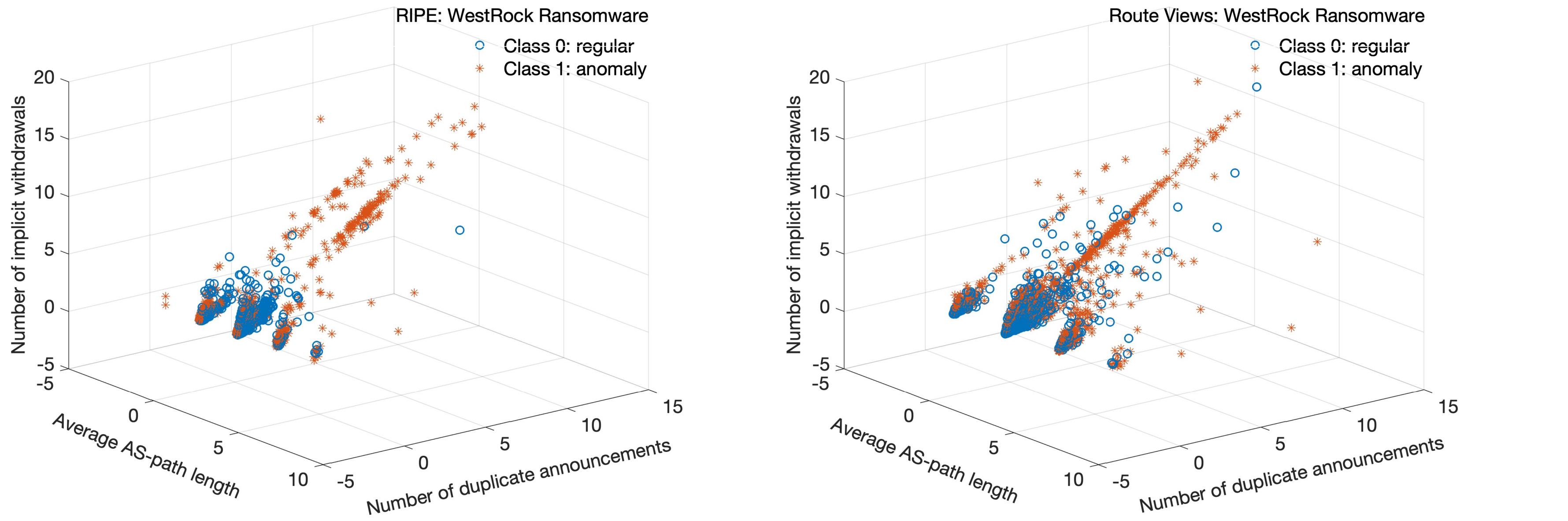
BGP Dataset: Pakistan Power Outage

- Number of announced NLRI prefixes vs. average unique AS-path vs. number of implicit withdrawals:



BGP Dataset: WestRock Ransomware Attack

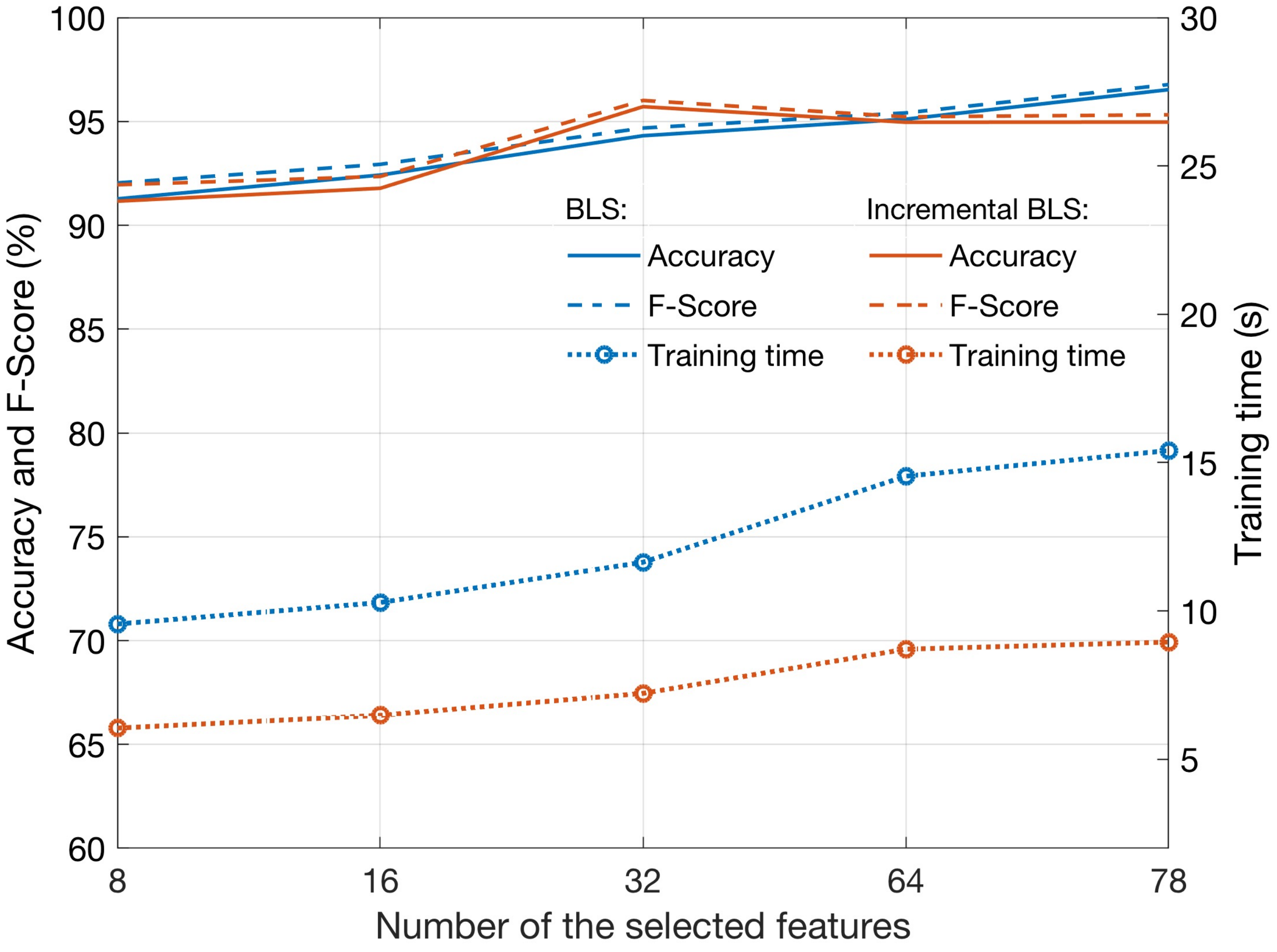
- Average unique AS-path vs. number of duplicate announcements vs. number of implicit withdrawals:



Roadmap

- Introduction
- Data processing:
  - BGP datasets
  - NSL-KDD dataset
- Machine learning models:
- Experimental procedure
- Performance evaluation
- Conclusions and references

Performance:
BLS and Incremental BLS, CIC 2017



Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
CNN	No refinement	RIPE	18.79	55.33	70.96	57.04	93.85
		RouteViews	18.66	57.67	72.96	58.04	98.23
	k-means	RIPE	19.31	55.31	71.00	56.99	94.25
		RouteViews	18.88	57.06	72.32	57.82	96.52
	Isolation forest	RIPE	19.20	55.29	70.96	57.00	94.00
		RouteViews	18.80	57.12	72.44	57.83	96.93

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
GRU_4	No refinement	RIPE	13.99	75.23	80.34	74.84	86.48
$LSTM_4$		RouteViews	18.95	55.42	70.72	57.20	92.60
GRU_4	k-means	RIPE	14.44	75.44	79.73	76.63	83.10
GRU_2		RouteViews	13.44	62.30	69.61	65.47	74.31
$LSTM_2$	Isolation forest	RIPE	12.63	75.36	79.73	76.41	83.35
$LSTM_3$		RouteViews	13.77	60.00	69.06	62.75	76.80

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
Bi-GRU$_4$	No refinement	RIPE	20.59	78.49	81.92	80.10	83.83
Bi-GRU$_3$		RouteViews	21.89	62.50	69.70	65.73	74.18
Bi-GRU$_3$	k-means	RIPE	20.27	77.76	82.05	77.30	87.43
		RouteViews	20.14	63.36	72.15	64.61	81.69
Bi-GRU$_4$	Isolation forest	RIPE	23.73	84.27	86.90	84.23	89.75
Bi-GRU$_3$		RouteViews	20.23	64.74	72.19	66.67	78.70

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
RBF-BLS	No refinement	RIPE	3.98	55.70	70.75	57.41	92.18
		RouteViews	2.60	54.74	69.99	56.95	90.78
	Isolation forest	RIPE	2.20	55.73	70.77	57.42	92.20
		RouteViews	3.97	54.61	69.81	56.91	90.28

Best Performance: WestRock Ransomware

Model Incremental		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
RBF-BLS	No refinement	RIPE	1.71	58.20	73.55	58.18	99.98
CEBLS		RouteViews	23.33	57.89	73.31	58.05	99.48
RBF-BLS	Isolation forest	RIPE	33.28	58.20	73.54	58.16	99.98
		RouteViews	7.01	58.15	73.52	58.16	99.93

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
VFBLS	No refinement	RIPE	7.31	55.15	70.18	57.19	90.80
		RouteViews	7.99	54.75	69.92	56.99	90.45
	Isolation forest	RIPE	6.18	54.74	69.81	57.00	90.05
		RouteViews	5.67	54.23	69.41	56.76	89.33

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
VCFBLS	No refinement	RIPE	4.14	55.33	70.31	57.30	90.95
		RouteViews	4.62	54.68	69.73	56.99	89.80
	Isolation forest	RIPE	6.56	54.72	69.86	56.98	90.27
		RouteViews	4.66	54.43	69.55	56.87	89.53

Best Performance: WestRock Ransomware

Model Incremental		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
VFBLS	No refinement	RIPE	6.77	58.17	73.54	58.16	100
		RouteViews	6.82	58.18	73.55	58.16	100
	Isolation forest	RIPE	11.60	58.27	73.55	58.23	99.80
		RouteViews	7.62	58.20	73.55	58.18	99.98

Best Performance: WestRock Ransomware

Model Incremental		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
VCFBLS	No refinement	RIPE	12.04	58.23	73.57	58.19	99.98
		RouteViews	9.08	58.30	73.57	58.25	99.85
	Isolation forest	RIPE	11.27	58.15	73.53	58.14	99.98
		RouteViews	10.40	58.20	73.56	58.17	100

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
XGBoost	No refinement	RIPE	0.54	60.44	73.38	60.26	93.80
		RouteViews	0.27	55.83	70.94	57.44	92.73
	Isolation forest	RIPE	0.52	59.84	73.05	59.88	93.62
		RouteViews	0.38	55.58	70.42	57.46	90.93

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
LightGBM	No refinement	RIPE	0.05	58.37	72.20	59.01	92.98
		RouteViews	0.06	57.50	72.16	58.27	94.73
	k-means	RIPE	0.14	58.11	71.29	59.25	89.48
		RouteViews	0.07	57.56	72.53	58.13	96.42
	Isolation forest	RIPE	0.10	57.66	71.42	58.77	91.02
		RouteViews	0.05	57.72	72.81	58.14	97.38

Best Performance: WestRock Ransomware

Model		Collection site	Training time (s)	Accuracy (%)	F-Score (%)	Precision (%)	Sensitivity (%)
CatBoost	No refinement	RIPE	0.33	55.60	71.36	57.09	95.15
		RouteViews	0.31	58.17	73.53	58.16	99.95
	Isolation forest	RIPE	0.32	55.58	71.34	57.07	95.12
		RouteViews	0.48	58.24	73.53	58.22	99.78

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

Conclusions

- We evaluated performance of:
  - CNNs
  - RNNs: LSTM, Bi-LSTM, GRU, and Bi-GRU deep recurrent neural networks with a variable number of hidden layers
  - BLS models with and without incremental learning:
    - radial basis function
    - cascades of mapped features and enhancement nodes
    - integrated extra-trees for feature selection (VFBLS and VCFBLS)
  - GBDT: XGBoost, LightGBM, CatBoost

Roadmap

- Introduction
- Traffic collection, characterization, and modeling
- Case studies: BCNET, E-Comm, ChinaSat, Internet
- Machine learning models
- Experimental procedure
- Performance evaluation
- Conclusions and references

References: Data Sources

- RIPE NCC:
  https://www.ripe.net
- University of Oregon Route Views project:
  http://www.routeviews.org
- NSL-KDD dataset:
  https://www.unb.ca/cic/datasets/nsl.html
- CICIDS2017 dataset:
  https://www.unb.ca/cic/datasets/ids-2017.html
- CSE-CIC-IDS2018 dataset:
  https://www.unb.ca/cic/datasets/ids-2018.html
- CAIDA: Center for Applied Internet Data Analysis:
  http://www.caida.org/home/

References: Tools

- Python: https://pypi.org

  Pandas: https://pandas.pydata.org/
- PyTorch
  https://pytorch.org/docs/stable/nn.html
- zebra-dump-parser:
  https://github.com/rfc1036/zebra-dump-parser
- BGP C# tool:
  http://www.sfu.ca/~ljilja/cnl/projects/BGP_datasets/index.html
- IEEE DataPort
  Border Gateway Protocol (BGP) datasets:
  - https://ieee-dataport.org/open-access/border-gateway-protocol-bgp-routing-records-reseaux-ip-europeens-ripe-and-bcnet
  - https://ieee-dataport.org/open-access/border-gateway-protocol-bgp-routing-records-route-views
- BLS: Broadlearning: http://www.broadlearning.ai/

References: Intrusion Detection

- J. P. A. Maranhão, J. P. C. L. da Costa, E. P. de Freitas, E. Javidi, and R. T. de Sousa, Jr., “Noise-robust multilayer perceptron architecture for distributed denial of service attack detection,” *IEEE Commun. Lett*., vol. 25, no. 2, pp. 402–406, Feb. 2021.
- P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, “A detailed investigation and analysis of using machine learning techniques for intrusion detection,” *IEEE Commun. Surveys Tut*., vol. 21, no. 1, pp. 686–728, First quarter 2019.
- A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” *IEEE Commun. Surv. Tut*., vol. 18, no. 2, pp. 1153–1176, 2016.
- M. C. Libicki, L. Ablon, and T. Webb, The Defenders Dilemma: Charting a Course Toward Cybersecurity. Santa Monica, CA, USA: RAND Corporation, 2015.
- V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: a survey,” *ACM Comput. Surv.,* vol. 41, no. 3, pp. 15:1–15:58, July 2009.

References: Deep Learning

- S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Oct. 1997.
- G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors,” *Computing Research Repository (CoRR)*, abs/1207.0580, pp. 1–18, Jul. 2012.
- K. Cho, B. van Merriënboer, C. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using RNN encoder–decoder for statistical machine translations,” in *Proc. 2014 Conf. Empirical Methods Natural Lang. Process*., Doha, Qatar, Oct. 2014, pp. 1724–1734.
- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning. Cambridge*, MA, USA: The MIT Press, 2016.
- K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, “LSTM: a search space odyssey,” *IEEE Trans. Neural Netw. Learn. Syst*., vol. 28, no. 10, pp. 2222–2232, Oct. 2017.

References: BLS and GBDT

- C. L. P. Chen and Z. Liu, “Broad learning system: an effective and efficient incremental learning system without the need for deep architecture,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.
- C. L. P. Chen, Z. Liu, and S. Feng, “Universal approximation capability of broad learning system and its structural variations,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 4, pp. 1191–1204, Apr. 2019.

- T. Chen and C. Guestrin, “XGBoost: a scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.
- G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, “LightGBM: a highly efficient gradient boosting decision tree,” in *Proc. Int. Conf. Neural Inform. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, 3146–3154.
- L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” in *Proc. Int. Conf. Neural Inform. Process. Syst.*, Montreal, Quebec, Canada, Dec. 2018, 6639–6649.

Publications: http://www.sfu.ca/~ljilja

Journal publication:

- Z. Li, A. L. Gonzalez Rios, and Lj. Trajkovic, “Machine learning for detecting the WestRock ransomware attack using BGP routing records,” *IEEE Communications* Magazine, vol. 61, no. 3, pp. 20–26, Mar. 2023.
- Z. Li, A. L. Gonzalez Rios, and Lj. Trajkovic, “Machine learning for detecting anomalies and intrusions in communication networks,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 2254-2264, July 2021.

Book chapters:

- Q. Ding, Z. Li, S. Haeri, and Lj. Trajković, “Application of machine learning techniques to detecting anomalies in communication networks: datasets and feature selection algorithms” in *Cyber Threat Intelligence,* M. Conti, A. Dehghantanha, and T. Dargahi, Eds., Berlin: Springer, pp. 47–70, 2018.
- Z. Li, Q. Ding, S. Haeri, and Lj. Trajković, “Application of machine learning techniques to detecting anomalies in communication networks: classification algorithms” in *Cyber Threat Intelligence,* M. Conti, A. Dehghantanha, and T. Dargahi, Eds., Berlin: Springer, pp. 71–92, 2018.

Publications: http://www.sfu.ca/~ljilja

Conference publications:

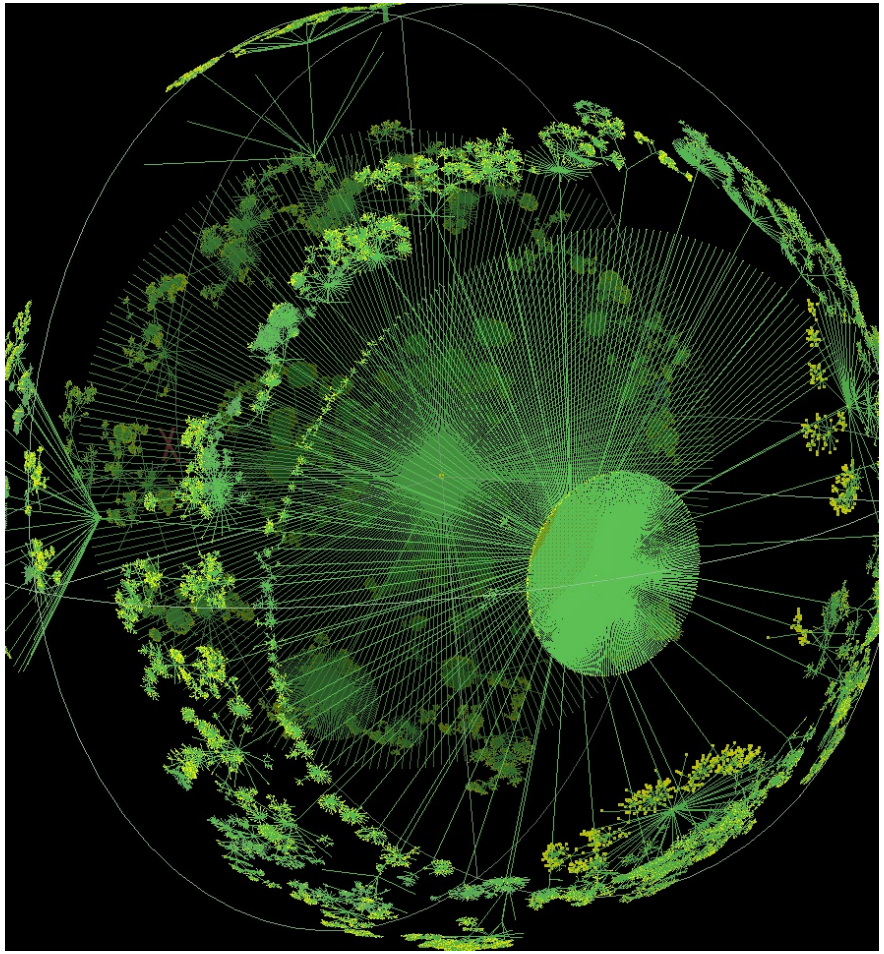
- Z. Li and Lj. Trajković, “CyberDefense: tool for detecting network anomalies and intrusions,” *IEEE Int. Conf. Syst., Man, Cybern.*, Honolulu, HI, USA, Oct. 2023.
- H. Takhar and Lj. Trajković, “BGP feature properties and classification of Internet worms and ransomware attacks,” *IEEE Int. Conf. Syst., Man, Cybern.*, Honolulu, HI, USA, Oct. 2023.
- T. Sharma, K. Patni, Z. Li, and Lj. Trajković, “Deep echo state networks for detecting Internet worm and ransomware attacks” *IEEE Int. Symp. Circuits and Systems*, Monterey, CA, USA, May 2023.
- Z. Li, A. L. Gonzalez Rios, and Lj. Trajković, “Classifying denial of service attacks using fast machine learning algorithms,” in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Melbourne, Australia, Oct. 2021, pp. 1221-1226 (virtual).
- K. Bekshentayeva and Lj. Trajkovic, "Detection of denial of service attacks using echo state networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Melbourne, Australia, Oct. 2021, pp. 1227-1232 (virtual).
- Z. Li, A. L. Gonzalez Rios, and Lj. Trajković, “Detecting Internet worms, ransomware, and blackouts using recurrent neural networks,” in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Toronto, Canada, Oct. 2020, pp. 2165-2172 (virtual).
- A. L. Gonzalez Rios, Z. Li, K. Bekshentayeva, and Lj. Trajković, “Detection of denial of service attacks in communication networks,” in *Proc. IEEE Int. Symp. Circuits and Systems*, Seville, Spain, Oct. 2020 (virtual).

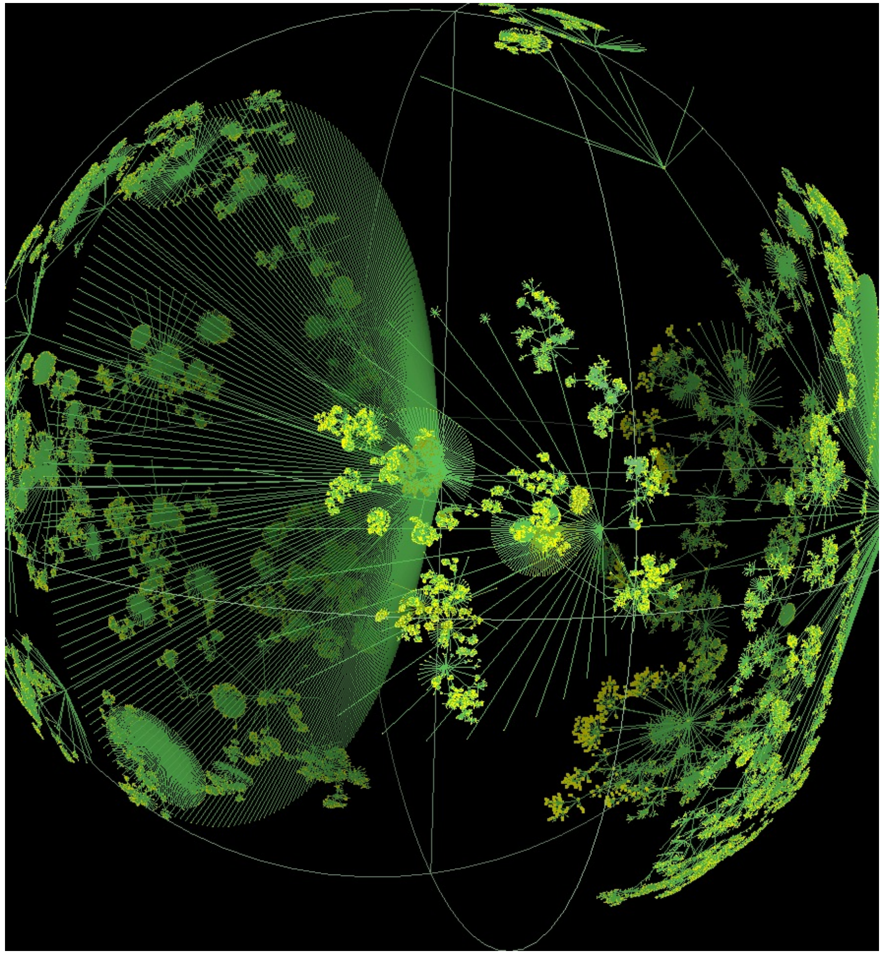
Publications: http://www.sfu.ca/~ljilja

Conference publications:

- Z. Li, A. L. Gonzalez Rios, G. Xu, and Lj. Trajković, "Machine learning techniques for classifying network anomalies and intrusions," in *Proc. IEEE Int. Symp. Circuits and Systems*, Sapporo, Japan, May 2019 (virtual).
- A. L. Gonzalez Rios, Z. Li, G. Xu, A. Dias Alonso, and Lj. Trajković, "Detecting network anomalies and intrusions in communication networks," in *Proc. 23rd IEEE International Conference on Intelligent Engineering Systems 2019*, Gödöllő, Hungary, Apr. 2019, pp. 29–34.
- Z. Li, P. Batta, and Lj. Trajković, "Comparison of machine learning algorithms for detection of network intrusions," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Miyazaki, Japan, Oct. 2018, pp. 4248–4253.
- P. Batta, M. Singh, Z. Li, Q. Ding, and Lj. Trajković, "Evaluation of support vector machine kernels for detecting network anomalies," in *Proc. IEEE Int. Symp. Circuits and Systems*, Florence, Italy, May 2018, pp. 1-4.
- Q. Ding, Z. Li, P. Batta, and Lj. Trajković, "Detecting BGP anomalies using machine learning techniques," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics,* Budapest, Hungary, Oct. 2016, pp. 3352–3355.

lhr: 535,102 nodes and 601,678 links





http://www.caida.org/home/