Detecting Domain Generation Algorithms Using Deep Learning

SRIRAM S

Center for Computational Engineering and Networking (CEN)

Amrita School of Engineering, Coimbatore

Amrita Vishwa Vidyapeetham

03rd June, 2019

Domain Name System (DNS)

The Domain Name System (DNS) is a critical component of the Internet infrastructure.

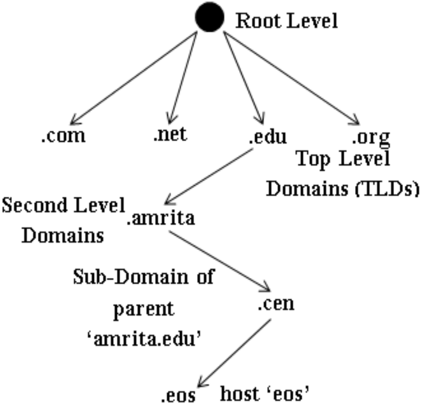

Figure 1: Hierarchical domain name system.

Domain Name System (DNS) (contd.)

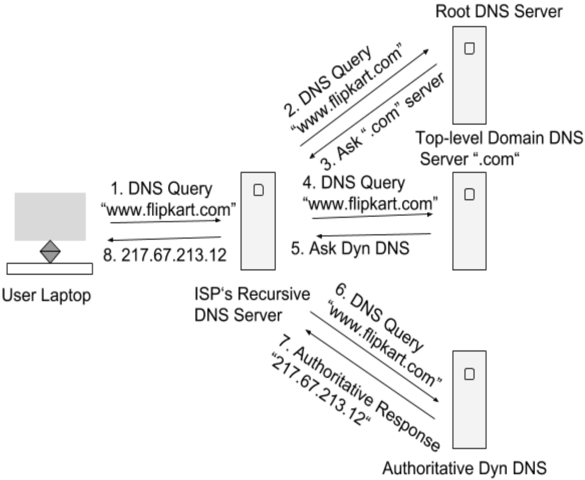

Figure 2: The DNS resolution process.

Domain Name System (DNS) (contd.)

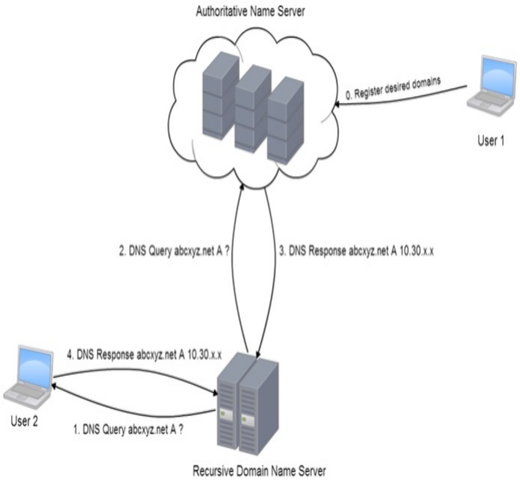

Figure 3: Working flow of a legitimate DNS query.

Domain Name System (DNS) (contd.)

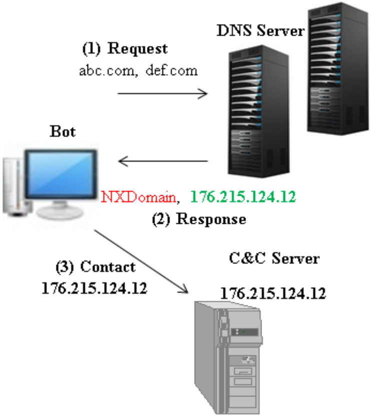

Figure 4: Domain-flux attacks.

Domain Generation Algorithms (DGAs) are popular: more than 70 DGAs known.

Domain Name System (DNS) (contd.)

- DGAs take a seed input and generate large amounts of pseudo-random domain names.
- A seed can be a date, a number, or any random characters.

```
def generate_domain(year, month, day, length=32, tld=''):
    """ Generates a domain by considering the current date. """
    domain = ""

    for i in range(length):
        year = ((year ^ 8 * year) >> 11) ^ ((year & 0xFFFFFFF0) << 17)
        month = ((month ^ 4 * month) >> 25) ^ 16 * (month & 0xFFFFFFF8)
        day = ((day ^ (day << 13)) >> 19) ^ ((day & 0xFFFFFFFE) << 12)
        domain += chr(((year ^ month ^ day) % 25) + 97)

    domain += tld

    return domain
```

```
btbpurnkbqidxxclfdfrdgjasjphyrtn.org
sehccrlyfadifehntnomqgpfyunqqfft.org
konsbolyfadifehntnomqpfyunqqfft.org
cytfiobnkjxomkhimxhcfvtogyaiqaa.org
```

Figure 5: CryptoLocker DGA.

Domain Name System (DNS) (contd.)

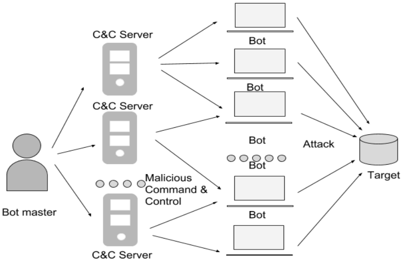

Figure 6: Botnet and Bot communication mechanism.

Identified problems is:

- Block the communication point between a bot and command and control (C2C) server using DNS data analysis.

Live stream DNS events collection in Ethernet LAN

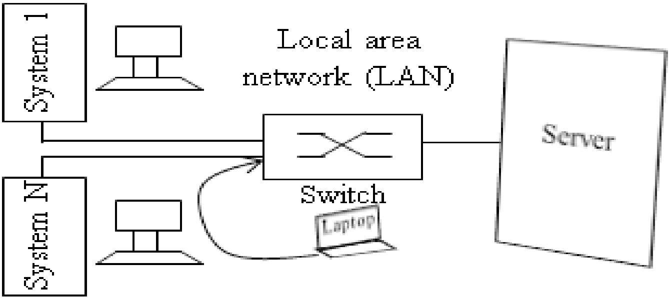

Figure 7: Port mirroring setup: duplicates traffic between different switch ports.

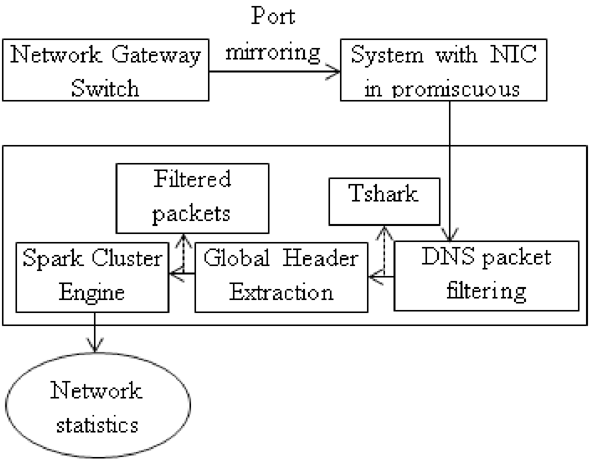

Figure 8: NIC in promiscuous mode.

Distributed DNS log parser

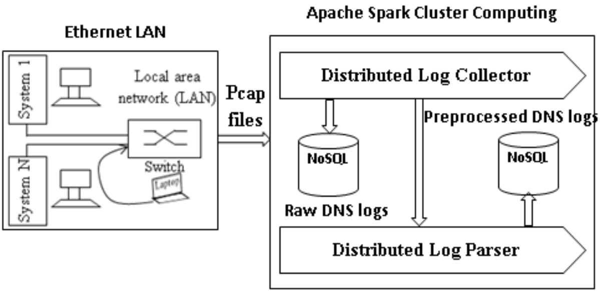

Figure 9: DNS data processing.

```
19:35:04.167395 IP censerver.local.27062 > 172.17.9.2.domain: 30578+ [b2&3=0x182] A?
www.mail.bel.co.in. (36)E..@..@.@.p...h...
.i..5.,c.wr...........www.mail.bel.co.in.....19:35:10.491014 IP censerver.local.65203 >
172.17.9.2.domain: 43048+ A? a.sitemeter.com. (33)E..=..@.@.p...h...    ....5.)t-.
(...........a   sitemeter.com.....19:35:10.491507 IP censerver.local.40442 >
172.17.9.2.domain: 42818+ A? www.google-analytics.com. (42)E..F..@.@.p...h...
....5.2>..B...........www.google-analytics.com.....19:35:11.387909 IP
censerver.local.61213 > 172.17.9.2.domain: 58471+ A? www.google.com. (32)
E..<..@.@.p...h...      ....5.(.L.g...........www.google.com.....19:35:11.402801 IP
censerver.local.32595 > 172.17.9.2.domain: 57996+ A? googleads.g.doubleclick.net. (45)
E..I..@.@.p...h...      ..5.5.5..............   googleads.g.doubleclick.net.....
19:35:11.402970 IP censerver.local.36159 > 172.17.9.2.domain: 1089+ A? r.casalemedia.com.
(35)E..?..@.@.p...h...  ..?.5.+*:.A...........r.casalemedia.com.....19:35:11.403070 IP
censerver.local.15131 > 172.17.9.2.domain: 18278+ A? t0.gstatic.com. (32)
E..<..@.@.p...h...      .;..5.(..Gf...........t0.gstatic.com.....19:35:11.403128 IP
censerver.local.65465 > 172.17.9.2.domain: 17500+ A? t3.gstatic.com. (32)
E..<..@.@.p...h...      ....5.(74D\...........t3.gstatic.com.....19:35:11.403248 IP
censerver.local.49894 > 172.17.9.2.domain: 60342+ A? www.facebook.com. (34)
E..>..@.@.p...h...      ....5.*...............www.facebook.com.....19:35:11.547008 IP
```

Figure 10: DNS log.

AmritaDGA Database

Table 1: Database statistics for classifying domain name into either legitimate or DGA.

Type	Legitimate	DGA generated
Training	655,683	135,056
Testing 1	2,349,331	108,076
Testing 2	182	2,740

Table 2: Database statistics for classifying domain name into either legitimate or DGA and categorizing DGA generated domain name to DGA family.

Family	Training	Testing 1	Testing 2
legitimate	100,000	120,000	40,000
banjori	15,000	25,000	10,000
corebot	15,000	25,000	10,000
dircrypt	15,000	25,000	300
dnschanger	15,000	25,000	10,000
fobber	15,000	25,000	800
murofet	15,000	16,667	5,000
necurs	12,777	20,445	6,200
newgoz	15,000	20,000	3,000
padcrypt	15,000	20,000	3,000
proslikefan	15,000	20,000	3,000
pykspa	15,000	25,000	2,000
qadars	15,000	25,000	2,300
qakbot	15,000	25,000	1,000
ramdo	15,000	25,000	800
ranbyus	15,000	25,000	500
simda	15,000	25,000	3,000
suppobox	15,000	20,000	1,000
symmi	15,000	25,000	500
tempedreve	15,000	25,000	100
tinba	15,000	25,000	700
Total	**397,777**	**587,112**	**103,200**

AmritaDGA Database Visualization

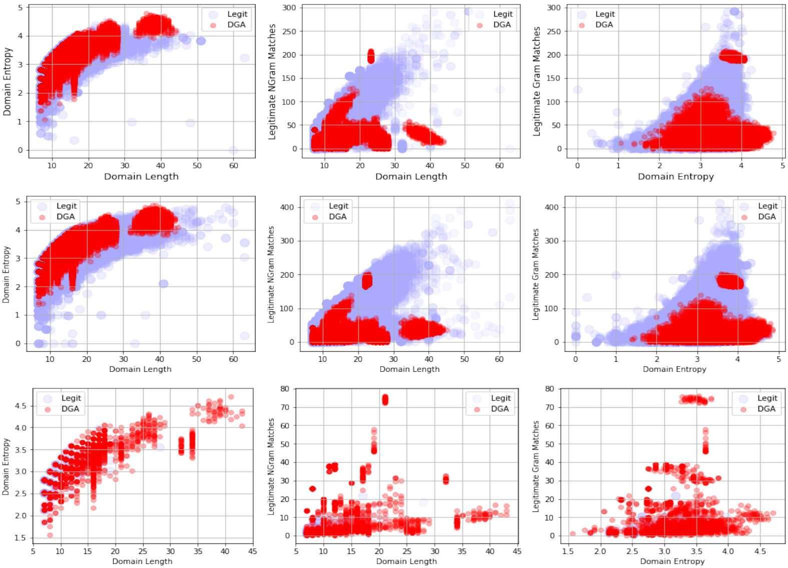

Figure 11: Training, Testing 1 and Testing 2 Visualization.

AmritaDGANet: Deep learning approach for DGA domain detection and classification

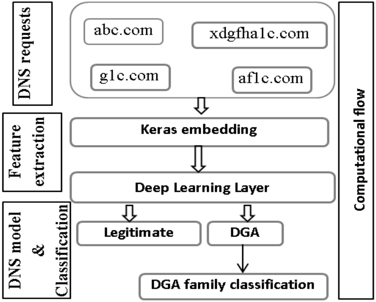

Figure 12: Work flow.

Hyperpamrameters: Embedding size: 128, Epochs: 100, Learning rate: 0.01, batch size: 64, optimizer: adam, No. of hidden layer: 1, No. hidden units: 128, and Dropout (only used in CNN): 0.04

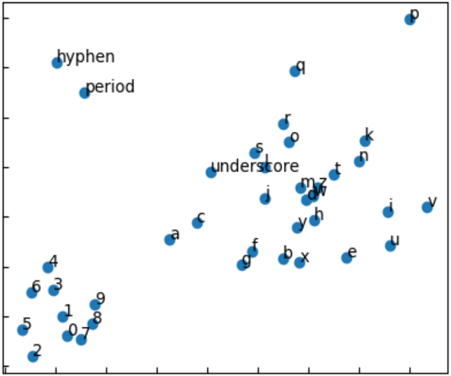

Figure 13: Character level embedded feature vectors learned by model are represented using two dimensional linear projection (PCA) with t-SNE. Note that models groups feature vectors based on the similarity.

Performance Evaluation

Table 3: Comparative Results of DGA domain detection and classification.

Model	Accuracy	Precision	Recall	F1-score
Binary classification				
RNN	97.9	68.8	94.4	79.6
	76.7	100	75.2	85.8
LSTM	98.8	79.7	96.0	87.1
	70.0	99.9	68.0	80.9
GRU	98.7	79.1	94.6	86.1
	71.8	99.9	70.0	82.3
CNN	97.8	67.3	96.5	79.3
	75.9	99.9	74.4	85.3
CNN-LSTM	98.5	77.2	93.8	93.8
	72.7	99.9	70.9	82.9
Multi-class classification				
RNN	66.2	62.7	66.2	60.9
	65.8	63.6	65.8	62.6
LSTM	66.9	69.5	66.9	62.7
	67.2	66.3	67.2	62.2
GRU	66.5	71.8	66.5	63.7
	64.9	65.5	64.9	60.1
CNN	64.3	69.1	64.3	59.6
	60.4	62.9	60.4	56.8
CNN-LSTM	65.8	67.6	65.8	62.5
	59.9	61.5	59.9	55.6

Shared task on detection of malicious domain names (DMD-2018) as part of SSCC'18 and ICACCI'18[1]. 19 teams registered, out of 19, 8 team submitted results and the paper. The dataset[2] and the baseline systems[3] are provided to the registered participants.

Table 4: DMD 2018 participated system results.

Team Name	Binary classification				Multi-class classification			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
UWT	**99**	96.6	82.8	89	63.3	61.8	63.3	60.2
	76.6	75.1	99.9	86	**88.7**	92.4	88.7	90.1
Deep_Dragons	98.7	95.5	78.7	86	**68.3**	68.3	68.3	64
	71.3	69.4	99.9	82	67.0	67.8	67	62.2
CHNMLRG	98.8	94.4	81.9	88	64.8	66.2	64.8	6
	78.7	77.4	99.9	87	**67.4**	68.3	67.4	64.8
BENHA	96.3	19.9	79.5	32	27.2	19.4	27.2	16.8
	56.4	55	97.4	70	42.9	34	42.9	27.2
BharathibSSNCSE	61.5	31.1	3.7	7	18	9.2	18	10.2
	56.2	55.9	95.6	71	33.5	22.9	33.5	22.3
UniPI	98.1	91.9	72.4	81	65.5	64.7	65.5	61.5
	71.4	69.6	99.9	82	67.1	64.1	67.1	61.9
Josan	**98.9**	94.7	82.2	88	**69.7**	68.9	69.7	65.8
	71.1	69.2	99.9	82	**67.9**	69.4	67.9	63.6
DeepDGANet	97.6	93.8	65.8	77	60.1	93.8	60.1	57.6
	78.2	76.9	99.7	87	53.1	65.3	53.1	54.1

[1] https://nlp.amrita.edu/DMD2018/
[2] https://vinayakumarr.github.io/AmritaDGA/
[3] https://github.com/vinayakumarr/DMD2018

After DMD 2018 shared task, the following institutions were given access:

- Ben-Gurion University, Beersheba, Israel.
- University of Washington, Tacoma, United states.
- CMC InfoSec Corp, VietNam, China.
- Akamai Technologies, United states.
- University of Murcia, Spain.
- Kansas State University, Manhattan, United States.
- University of Science and Technology, Algeria.
- Georgia Institute of Technology, Atlanta, Georgia, United states.
- Graduate School of Information Security, Korea University
- Vellore Institute of Technology, Chennai, India.
- IIT Kanpur, India.
- Xidian University, China.
- University of Pisa.
- Mangalore University.
- PES University, India.
- Savitribai Phule Pune University.
- Punjabi University, Patiala.
- SSN College of Engineering, Coimbatore.

Large-scale Learning: Improved DGA detection

Table 5: Results of RNN- classical machine learning algorithms (CMLAs).

Method	Testing 1				Testing 2			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
RNN - LR	**66.5**	64.5	66.5	63.1	66.5	62.6	66.5	60.8
RNN - NB	57.3	59.5	57.3	54.8	63.0	68.9	63.0	62.2
RNN - KNN	65.8	62.8	65.8	62.2	66.4	64.0	66.4	61.1
RNN - DT	61.3	60.4	61.3	58.8	63.9	632	63.9	59.3
RNN - RF	65.4	62.7	65.4	61.8	66.4	634	66.4	60.7
RNN - SVM-L	**66.2**	63.2	66.2	62.4	66.4	63.3	66.4	60.4
RNN - SVM-RBF	67.0	63.5	67.0	63.1	66.7	62.8	66.7	61.0

LR: Logistic regression, NB: Naive Bayes, KNN: K-nearest neighbour, DT: Decision tree, RF: Random forest, SVM-L: Support vector machine with linear kernel and SVM-RBF: Support vector machine with RBF kernel.

Table 6: Results of LSTM- classical machine learning algorithms (CMLAs).

Method	Testing 1				Testing 2			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
LSTM - LR	**67.4**	67.4	67.4	63.2	66.9	69.9	66.9	63.0
LSTM - NB	60.8	61.9	60.8	57.1	64.2	67.0	64.2	61.9
LSTM - KNN	66.6	65.4	66.6	62.0	66.5	68.0	66.5	62.6
LSTM - DT	62.8	63.1	62.8	59.2	64.6	67.0	64.6	61.4
LSTM - RF	65.6	66.3	65.6	60.9	66.5	67.2	66.5	62.2
LSTM - SVM-L	67.1	66.4	67.1	62.5	66.8	70.1	66.8	62.9
LSTM - SVM-RBF	66.8	66.0	66.8	61.8	66.8	67.2	66.8	62.7

Table 7: Results of GRU- classical machine learning algorithms (CMLAs).

Method	Testing 1				Testing 2			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
GRU - LR	**65.1**	64.3	65.1	59.6	**66.6**	70.7	66.6	63.3
GRU - NB	58.4	62.9	58.4	54.9	62.1	69.6	62.1	61.1
GRU - KNN	**65.3**	63.2	65.3	60.5	**66.6**	70.6	66.6	63.4
GRU - DT	60.8	60.3	60.8	56.7	64.1	68.2	64.1	61.9
GRU - RF	**64.5**	63.0	64.5	59.0	66.2	68.9	66.2	62.7
GRU - SVM-L	**65.0**	65.1	65.0	59.2	**66.7**	710	66.7	63.2
GRU - SVM-RBF	**65.2**	64.5	65.2	59.4	**66.5**	68.5	66.5	62.9

Table 8: Results of CNN- classical machine learning algorithms (CMLAs).

Method	Testing 1				Testing 2			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
CNN - LR	59.0	62.3	59.0	55.3	62.7	65.2	62.7	59.6
CNN - NB	53.0	58.1	53.0	50.4	58.2	61.7	58.2	55.1
CNN - KNN	**60.7**	62.4	60.7	58.1	62.0	64.9	62.0	58.7
CNN - DT	55.7	58.2	55.7	51.3	59.8	63.4	59.8	56.3
CNN - RF	59.4	61.1	59.4	54.4	62.9	63.1	62.9	58.2
CNN - SVM-L	58.2	57.3	58.2	52.9	63.6	61.7	63.6	58.6
CNN - SVM-RBF	20.6	19.5	20.6	7	38.8	20.1	38.8	21.8

Table 9: Results of CNN-LSTM- classical machine learning algorithms (CMLAs).

Method	Testing 1				Testing 2			
	Accuracy	**Precision**	**Recall**	**F1-score**	**Accuracy**	**Precision**	**Recall**	**F1-score**
CNN-LSTM - LR	59.6	61.9	59.6	55.8	65.3	69.0	65.3	62.7
CNN-LSTM - NB	53.6	53.9	53.6	49.9	61.4	66.6	61.4	59.7
CNN-LSTM - KNN	59.3	60.5	59.3	55.7	64.9	69.7	64.9	62.4
CNN-LSTM - DT	55.2	57.2	55.2	51.7	61.9	65.2	61.9	59.4
CNN-LSTM - RF	58.4	60.8	58.4	53.8	64.8	65.1	64.8	60.6
CNN-LSTM - SVM-L	59.2	61.3	59.2	54.6	65.6	68.9	65.6	62.4
CNN-LSTM - SVM-RBF	59.3	61.6	59.3	54.6	65.3	68.6	65.3	62.1

Performance enhancement using character level deep learning architectures

Table 10: Character level deep learning architectures.

Name	Architecture	Task
Endgame, (Woodbridge et al, 2016)	LSTM	Detecting and categorizing domain names that are generated by DGAs
Invincea, (Saxe et al, 2017)	CNN	To detect malicious URLs, file paths and registry keys
CMU, (Dhingra et al, 2016)	Bidirectional recurrent structures	Social media text classification, Twitter
MIT, (Vosoughi et al, 2016)	Hybrid of CNN and LSTM	Social media text classification, Twitter
NYU, (Zhang et al, 2015)	Stacked CNN layers	Text classification

Table 11: Results of character level deep learning architectures.

Model	Accuracy	Precision	Recall	F1-score
Binary classification				
Endgame	99.2	85.2	99.2	**91.7**
	80.7	99.9	79.5	**88.5**
Invincea	99.2	84.9	99.2	**91.5**
	79.6	99.9	78.3	**87.8**
CMU	99.2	85.2	99.2	**91.7**
	82.0	99.9	80.9	**89.4**
MIT	99.2	85.1	99.2	**91.6**
	81.5	99.9	80.3	**89.1**
NYU	99.2	85.1	99.2	**91.6**
	80.1	99.9	78.9	**88.2**
Multi-class classification				
CMU	71.2	69.7	67.1	**68.4**
	89.1	93.1	90.1	**91.5**

Cost-sensitive deep learning architecture to handle multi-class imbalanced problem

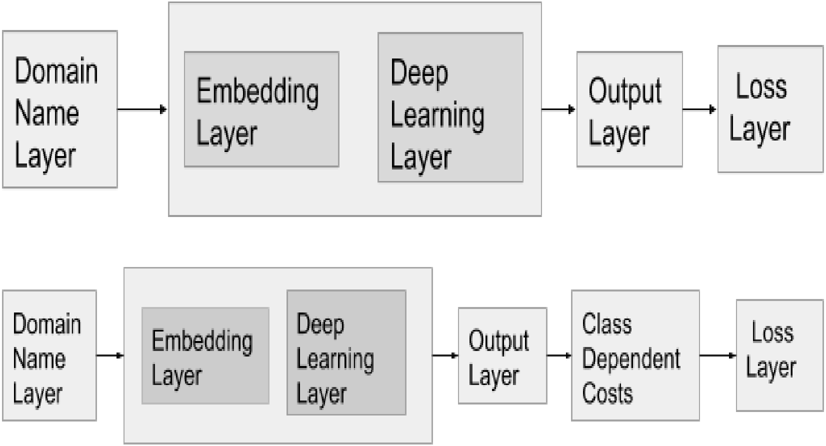

Figure 14: Architecture of cost-insensitive and cost-sensitive deep learning architecture, all connections are not shown.

Table 12: Results of character level cost-sensitive deep learning architectures.

Model	Accuracy	Precision	Recall	F1-score
Binary classification				
C-Endgame	99.2	85.2	99.2	91.7
C-Endgame	83.8	99.9	82.8	**90.6**
C-Invincea	99.2	85.1	99.2	**91.6**
C-Invincea	82.5	99.9	81.5	**89.7**
C-CMU	99.2	85.4	99.2	**91.8**
C-CMU	84.5	99.9	83.5	**91.0**
C-MIT	99.2	85.3	99.2	**91.7**
C-MIT	84.1	99.9	83.1	**90.7**
C-NYU	99.2	85.0	99.2	**91.5**
C-NYU	83.2	99.9	82.2	**90.2**
Multi-class classification				
C-CMU	73.1	72.8	70.1	**71.4**
C-CMU	89.9	93.4	90.5	**91.9**

Domain name spoofing

Domain name spoofing approach creates domain names that are visually similar to legitimate and recognized names.

Domain name			Type
netflixlife.com	**instagram.com**	**alibaba.com**	**Legitimate**
netflixlifel.com	instagra44.com	al1baba.com	Homoglyph
ne_vflixlife.com	hnstagzam.com	alibba.com	Homoglyph
nevflixnifem.com	insfagza_m.com	aia6ba.com	Homoglyph
netflixlfe.com	nstagr4m.com	al_ibaba.com	Homoglyph

Table 13: The first row is the legitimate domain name and other four rows are homoglyph attacks.

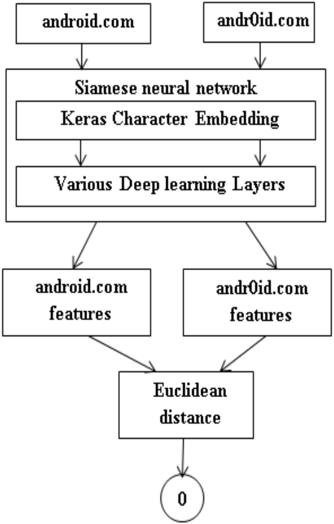

Figure 15: Domain name Similarity checker using Siamese neural network.

Performance evaluation

Table 14: Statistics of Domain name dataset.

Type	#Samples	
	Similar	Dissimilar
Train	348,615	627,507
Validation	18,350	33,030
Test	91,745	165,141

Table 15: Statistics of Process name dataset.

Type	#Samples	
	Similar	Dissimilar
Train	413,124	677,864
Validation	103,281	35,669
Test	129,102	178,419

Both databases are obtained from (Woodbridge et al, 2018).

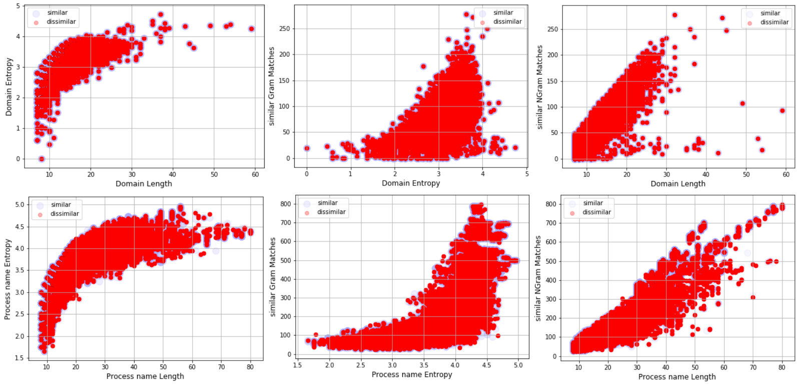

Figure 16: Domain name and Process name data Visualization.

Table 16: Performance of Siamese networks in terms of Receiver operating characteristic - Area under curve (ROC-AUC).

Method	ROC-AUC	
	Domain Name Spoofing	Process Name Spoofing
S-CNN (Woodbridge et al, 2018)	0.96	0.80
S-RNN (Proposed)	0.78	0.75
S-IRNN (Proposed)	0.96	0.70
S-LSTM (Proposed)	**0.97**	**0.97**
S-GRU (Proposed)	**0.97**	0.96
S-B-RNN (Proposed)	**0.97**	0.93
S-B-IRNN (Proposed)	0.80	0.77
S-B-LSTM (Proposed)	**0.97**	0.96
S-B-GRU (Proposed)	0.96	0.95
VED (Woodbridge et al, 2018)	0.89	0.43
ED (Woodbridge et al, 2018)	0.81	0.51
PED (Woodbridge et al, 2018)	0.86	0.44

Table 17: Parameter details of Siamese networks.

Method	Domain Name Spoofing	Process Name Spoofing
	Parameters	Parameters
S-CNN (Woodbridge et al, 2018)	148,832	148,832
S-RNN (Proposed)	58,496	58,496
S-IRNN (Proposed)	58,496	58,496
S-LSTM (Proposed)	**157,184**	**157,184**
S-GRU (Proposed)	**124,288**	124,288
S-B-RNN (Proposed)	**91,392**	91,392
S-B-IRNN (Proposed)	104,192	104,192
S-B-LSTM (Proposed)	**288,768**	288,768
S-B-GRU (Proposed)	222,976	222,976

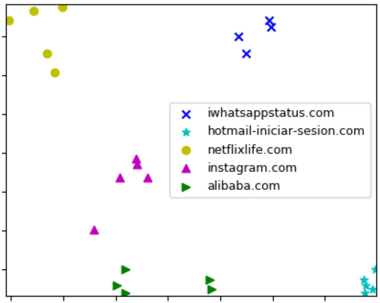

Figure 17: t-SNE visualization.

Table 18: The first row is the legitimate domain name and other four rows are spoofed domain names.

Domain name					Type
iwhatsappstatus.com	**hotmail-iniciar-sesion.com**	**netflixlife.com**	**instagram.com**	**alibaba.com**	**Legitimate**
iwhatsappstadus.com	hotmail-iniciar-serion.com	netflixlifel.com	instagra44.com	al1baba.com	Homoglyph
iwhatsappsfatuw.com	hotiail-inigiar-sesion.com	ne_vflixlife.com	hnstagzam.com	alibba.com	Homoglyph
iwhatsapps-tatu_w.com	hottiaih-iniciar-sesion.com	nevflixnifem.com	insfagza_m.com	aia6ba.com	Homoglyph
iwhatsapstatus.com	hotmail-inicar-sesion.com	netflixlfe.com	nstagr4m.com	al_ibaba.com	Homoglyph

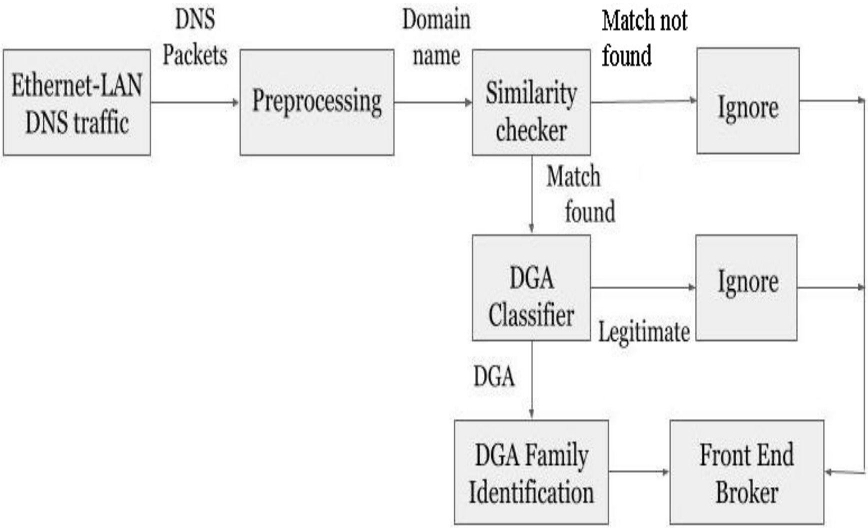

Figure 18: Two-Level Framework for Domain Name Systems Data Analysis.

Contributions of the present work

The major contributions are:

- Developed generated domain-flux attacks database for anomaly intrusion detection systems.
- Proposed a novel and unified deep learning based two-level framework for DNS data analysis in the Ethernet level.

Limitations for the present work and Scope for the future work

- DGA detection: Embedding representation is specific to the training data and is not representative of English language. This type of embedding can improve detection accuracy for unknown DGA malware.
- Multi-lingual Internationalized Domain Names (IDN) domain name support.

References

[1] Ashok, A., Poornachandran, P., Pal, S., Sankar, P., & Surendran, K. (2017). Why so abnormal? Detecting domains receiving anomalous surge traffic in a monitored network. Journal of Intelligent & Fuzzy Systems, 32(4), 2901-2907.

[2] Antonakakis, M., Perdisci, R., Nadji, Y., Vasiloglou, N., Abu-Nimeh, S., Lee, W., & Dagon, D. (2012). From throw-away traffic to bots: detecting the rise of DGA-based malware. In Presented as part of the 21st USENIX Security Symposium (USENIX Security 12) (pp. 491-506).

[3] Anderson, H. S., Woodbridge, J., & Filar, B. (2016, October). DeepDGA: Adversarially-tuned domain generation and detection. In Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security (pp. 13-21). ACM.

[4] Antonakakis, M., Perdisci, R., Lee, W., Vasiloglou, N., & Dagon, D. (2011, August). Detecting Malware Domains at the Upper DNS Hierarchy. In USENIX security symposium (Vol. 11, pp. 1-16).

[5] J. Woodbridge, H. S. Anderson, A. Ahuja, and D. Grant, Predicting domain generation algorithms with long short-term memory networks, preprint arXiv:1611.00791, 2016.

[6] Woodbridge, J., Anderson, H. S., Ahuja, A., & Grant, D. (2018, May). Detecting Homoglyph Attacks with a Siamese Neural Network. In 2018 IEEE Security and Privacy Workshops (SPW) (pp. 22-28). IEEE.

[7] Sun, Y., Kamel, M. S., Wong, A. K., & Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition, 40(12), 3358-3378.

Thank You . . .