Сравнительная характеристика подходов к анализу трафика сети даркнет с использованием методов машинного обучения

Д.А. Левшун (gaifulina@comsec.spb.ru)

Обозначения и сокращения:

Задача:

- ИАТ – идентификация анонимного трафика,
- КП – классификация типов/протоколов трафика,
- КС – классификация сервисов трафика,
- ОВП – обнаружение вредоносных программ.

Базовая модель:

- DT – дерево решений (Decision Tree),
- RF – случайный лес (Random Forest),
- SVM – метод опорных векторов (Support Vector Machine),
- KNN – метод k-ближайших соседей (K-Nearest Neighbors),
- LR – линейная регрессия (Linear Regression),
- NB – наивный Байесовский классификатор (Naïve Bayes),
- GB – градиентный бустинг (Gradient Boosting,)
- AdaBoost – адаптивный бустинг (Adaptive Boosting),
- ANN – искусственная нейронная сеть (Artificial Neural Network),
- DBSCAN – основанная на плотности пространственная кластеризация для приложений с шумами (Density-Based Spatial Clustering of Applications with Noise),
- MLP – многослойный перцептрон (MultiLayer Perceptron),
- CNN – сверточная нейронная сеть (Convolutional Neural Network),
- TCN – временная сверточная сеть (Temporal Convolutional Network),
- LSTM – блок долгой краткосрочной памяти (Long Short-Term Memory),
- BiLSTM – двунаправленный блок долгой краткосрочной памяти (Bidirectional Long Short-Term Memory),
- GRU – управляемый рекуррентный блоках (Gated Recurrent Unit),
- Bi GRU – двунаправленный управляемый рекуррентный блоках (Bidirectional Gated Recurrent Unit),
- GAN – генеративно-состязательная сеть (Generative Adversarial Network,),
- AE – автокодировщик (AuroEncoder),
- SAE – разреженный автокодировщик (Sparse AuroEncoder),
- GCN – графовые сверточные сети (Graph Convolutional Network).

Публикация	Тип трафика	Задача	Базовая модель	Набор данных
Abu Al-Haija et al. (2022)	Tor, VPN	КП	DT, AdaBoost, KNN	CIC-Darknet2020
Akshobhya (2021)	Tor, I2P, JonDonym	КП	KNN, NB, DT, RF	Anon17
Alimoradi et al. (2022)	Tor, VPN	КП	MLP	CIC-Darknet2020
Almomani (2023)	Tor	ИАТ	SVM, RF, LR, MLP	CIC-Darknet2020
Anil and Shina (2020)	VPN	ИАТ, КП	Mean Shift	Собственный
Bader et al. (2022)	VPN	КС, ОВП	CNN, GRU, LSTM	Собственный
Bakhsh et al. (2023)	Tor, VPN	ИАТ	MLP, LSTM	CIC-Darknet2020
Bar and Hajaj (2022)	VPN	ИАТ, ОВП	Word2Vec	ISCXVPN2016, USTC-TFC2016
Cai et al. (2019)	Tor, I2P, JonDonym	КС	RF, GB	Anon17
Choorod et al. (2024)	Tor	КП, КС	DT (J48), RF, KNN	ISCXTor2016
Choorod and Weir (2021)	Tor	КС	DT (J48), KNN, RF	CIC-Darknet2020
Cohen et al. (2020)	VPN	ИАТ	Word2vec, DBSCAN	Собственный
Cong Dong et al. (2020)	VPN	КС	KNN, RF, GB, CNN, AE	ISCX VPN2016
Demertzis et al. (2021)	Tor, VPN	ИАТ	Weight-Agnostic Neural Network	CIC-Darknet2020
Dodia et al. (2022)	Tor	ОВП	KNN, LR, GB, RF	VirusTotal
Gioacchini et al. (2023)	VPN	ИАТ, ОВП	Word2Vec, KNN	Собственный
Guan et al. (2019)	Tor	КП	DT, NB, LR, RF, GB	CIC-Darknet2020
Gudla et al. (2024)	Tor	КС	DT, SVM, GB, MLP	ISCXTor2016
Han et al. (2022)	VPN	ОВП	Graphical Lasso	Собственный
Hardhik et al. (2022)	Tor, VPN	ИАТ	DT, KNN, RF	CIC-Darknet2020
He et al. (2023)	Tor	КС	DT (J48), RF, NB, MLP, CNN	ISCXTor2016
Hou et al. (2024)	Tor, VPN	КС	MLP	CIC-Darknet2020
Hu et al. (2020)	Tor, I2P, Zeronet, Freenet	КП	LR, DT, RF, GB, MLP, LSTM	DarknetDataset-2020
Hurali et al. (2020)	Tor, I2P, JonDonym	КП, КС	KNN, SVM, DT, RF	Anon17
Hwang et al. (2019)	VPN	ОВП	CNN, LSTM	USTC-TFC2016
Iliadis et al. (2021)	Tor, VPN	КП	KNN, MLP, RF, DT, GB	CIC-Darknet2020
Ishikawa et al. (2020)	VPN	ОВП	FastText, DBSCAN	Собственный
Islam et al. (2023)	Tor, VPN	КС	CNN, SAE	ISCXTor2016, ISCXVPN2016
Karunanayake et al. (2023)	Tor	КС	KNN, RF, SVM	Tor dataset
Khajehpour et al. (2022)	Tor	ОВП	MLP, CNN, LSTM	Собственный
Kim and Anpalagan (2018)	Tor	КП	CNN	ISCXTor2016
Kumar et al. (2019)	VPN	ИАТ	Microsoft Azure ML	Собственный
Lan et al. (2022)	Tor, VPN	КС	CNN, BiLSTM	CIC-Darknet2020
Lashkari et al. (2020)	Tor, VPN	КС	CNN	ISCXTor2016, ISCXVPN2016
Li and Xu (2024)	VPN	ИАТ, ОВП	CNN, TCN	ISCXVPN2016, USTC-TFC2016
Li et al. (2021)	VPN	КС	CNN	ISCXVPN2016, USTC-TFC2016

Lin K. et al. (2021)	Tor	КП, КС	CNN, LSTM	ISCXTor2016
Lin X. et al. (2022)	Tor, VPN, TLS	КС	BERT	CSTNET-TLS, ISCXTor2016, ISCXVPN2016, USTC-TFC2016
Lotfollahi et al. (2020)	VPN	КП, КС	SAE, CNN	ISCXVPN2016
Luo et al. (2023)	VPN	КС	KNN, SVM, NB, RF, GB, MLP	ISCXVPN2016, USTC-TFC2016
Marim et al. (2023)	Tor, VPN	ИАТ, КС	DT, RF, MLP	CIC-Darknet2020
Mohanty et al. (2022)	Tor, VPN	ИАТ	DT, KNN, RF, AE	CIC-Darknet2020
Montieriet al. (2020)	Tor, I2P, JonDonym	КС	DT (C4.5), NB, RF	Anon17
Niu et al. (2024)	Tor, VPN	КС	CNN, LSTM	ISCXTor2016, ISCXVPN2016
Petagna et al. (2019)	Tor	КС	SVN, KNN, RF	Peel the Onion
Rust-Nguyen et al. (2023)	Tor, VPN	ИАТ	SVM, RF, GB, KNN, MLP, CNN, GAN	CIC-Darknet2020
Sarkar et al. (2020)	Tor	КП	MLP	ISCXTor2016
Sarwar et al. (2021)	Tor, VPN	КП, КС	DT, GB, RF, CNN, LSTM	CIC-Darknet2020
Shapira et al. (2019)	Tor, VPN	КС	CNN	ISCXTor2016, ISCXVPN2016
Shi K. et al. (2022)	Freenet	ИАТ, КП	KNN	DarknetDataset-2020
Shi Z. et al. (2023)	VPN	КС	BERT, CNN	ISCXVPN2016
Singh et al. (2021)	Tor, VPN	ИАТ	CNN, SVM, DT, RF	ISCXTor2016, ISCXVPN2016
Sridhar and Sanagavarapu (2021)	Tor	КП	RF, GAN	CIC-Darknet2020
Sun et al. (2022)	Tor, VPN	КС	CNN	ISCXTor2016
Vishnupriya et al. (2023)	Tor	КП	LSTM	ISCXTor2016
Wang B. et al. (2020)	VPN	ОВП	CNN, GRU	USTC-TFC2016
Wang L. et al. (2020)	Tor	КП, КС	DT (J48), NB	Собственный
Xu et al. (2023)	Tor, VPN	КС	MLP	ISCXVPN2016, USTC-TFC2016
Yang J. et al. (2023)	Tor, VPN	ИАТ	Transformer, BiLSTM	CIC-Darknet2020
Yang T. et al. (2023)	Tor, VPN	КП	AE	CIC-Darknet2020
Yao et al. (2019)	VPN	КС	BiLSTM	ISCX VPN2016
Yin et al. (2022)	Tor	КС	CNN	ISCXTor2016
Zeng et al. (2019)	VPN	КС	CNN, LSTM, SAE	ISCX VPN2016
Zhai et al. (2023)	Tor, VPN	КС	CNN, BiGRU	CIC-Darknet2020
Zhao R. et al. (2021)	Tor, I2P, JonDonym	КС	CNN, LSTM	Anon17, ISCXVPN2016, SJTU-AN21
Zhao R. et al. (2022)	Tor, I2P, JonDonym	КС	GCN	Anon17
Zheng et al. (2020)	VPN	КС	AE	ISCX VPN2016
Буковшин и соавт. (2020)	Tor, VPN	КС	AE	ISCXTor2016, ISCXVPN2016
Мулюха и соавт. (2020)	VPN	КС	NB	Собственный
Старун и Югасон (2022)	VPN	КП	GB	ISCXVPN2016