SINCONF 2020
13th International Conference on
Security of Information and Networks
4-7 November 2020
Istanbul, Turkey.

Neural network approach to assessing cybersecurity risks in large-scale dynamic networks

Vasiliy
KRUNDYSHEV

vmk@ibks.spbstu.ru

POLYTECH
Peter the Great
St. Petersburg Polytechnic
University

DEPARTMENT
Information Security
of Computer Systems

The reported study was funded by RFBR according to the research project № 19-37-90001.

Relevance of work

- According to experts, by 2024 the number of geographically distributed networks of smart cities will be about 1.3 billion.
- According to the estimates of the research company ABI Research, $ 135 billion will be spent on cybersecurity of the vital infrastructure of smart cities by 2024

- the main threats are those aimed at **disrupting the management processes of enterprises and urban infrastructure**, not at stealing personal data.

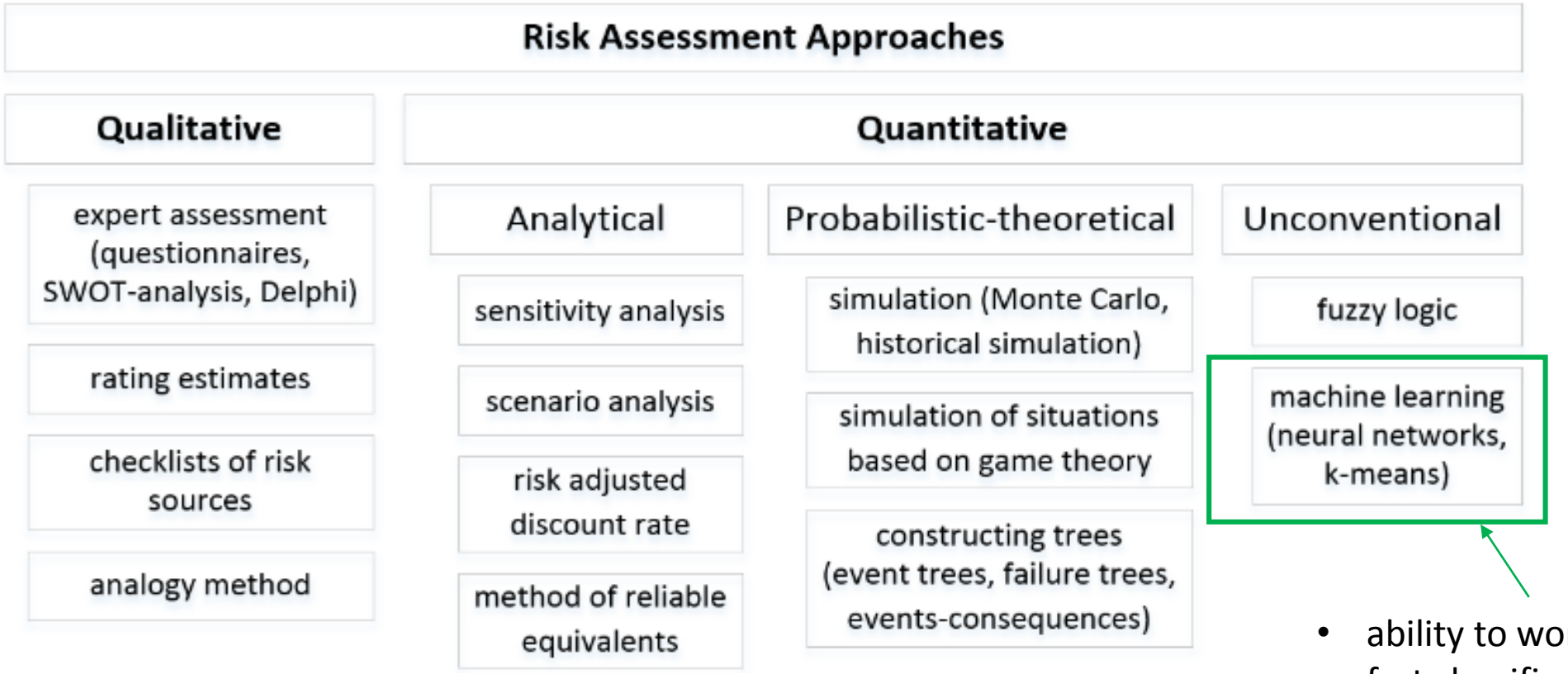
Goal and tasks

The goal is to design a method for dynamic assessment of cyber risks, taking into account the features of the smart city ecosystem.

Tasks:

1. Review and analysis of the related works for risk assessment .
2. Discussing our method for assessment of cyber risks.
3. Experimental modeling and testing of the developed method.

Classification of known approaches to risk assessment



- ability to work with big data
- fast classification speed
- discovering hidden patterns
- high accuracy

Existing methods of risk analysis and our requirements

ПОЛИТЕХ
Санкт-Петербургский политехнический университет Петра Великого

Methods / requirements	Quantitative approach	Speed computing	Large number of devices	Interaction and influence	Versatility
Microsoft	+/-	-	+	-	+
CRAMM	+/-	-	+	-	+
CORAS	+/-	-	-	-	-
OCTAVE	-	-	-	-	-
FRAP	-	-	-	-	-
GRIF	+	+	+	-	+
RiskWatch	+	+	+	-	+

Analysis of existing methods of security assessment

Risk assessment methods based on expert assessments and requiring the active participation of a person, cannot be applied in dynamic infrastructures.

Methods that use scenario analysis and functional analysis are industry-specific, and poorly adaptable to address cybersecurity challenges.

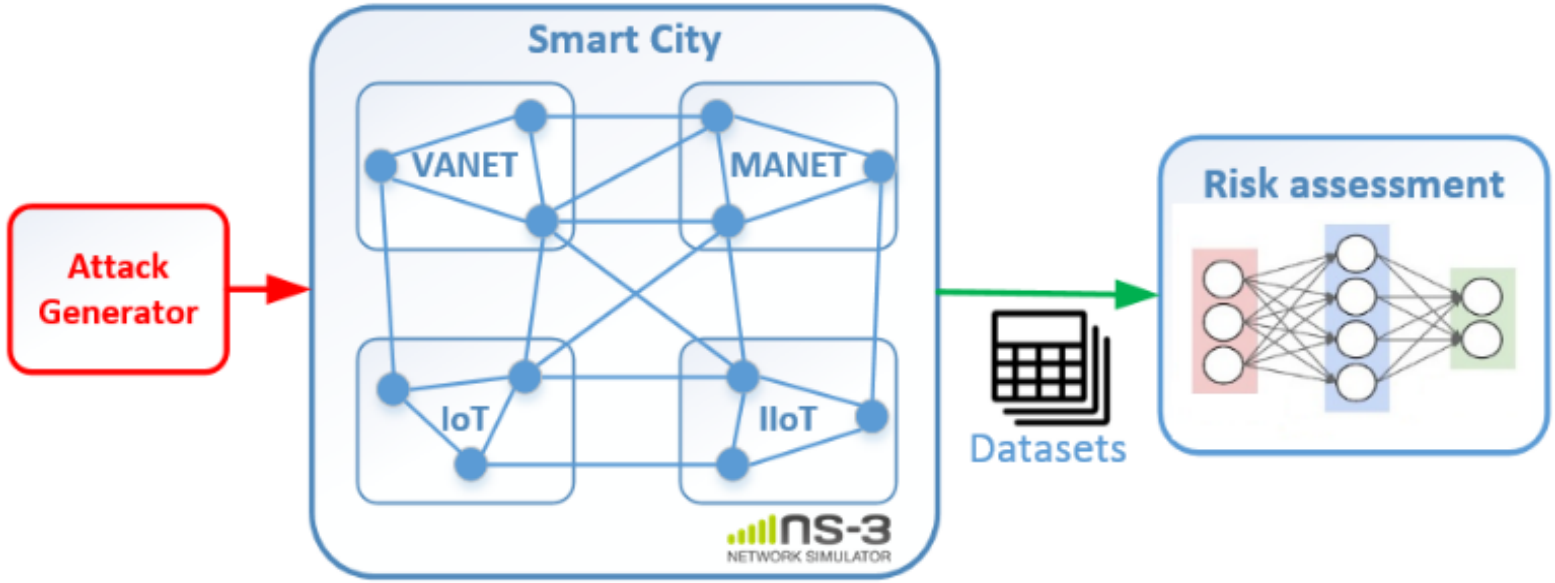
Statistical methods is hampered by the complexity of collecting statistical data for modeling calculations of the resulting indicators in networks with a peer-to-peer architecture, as well as the dependence of accuracy of decisions on the number of iterations.

Methods based on **artificial intelligence**, due to their adaptability and predictability, are most suitable for assessing cybersecurity risks in dynamic networks of a smart city. The quantitative approach used in neural networks sets the exact values of the probability of threats and possible consequences, as well as the risk itself for each type of asset. Numerical values are convenient for analyzing and comparing results.

Experimental setup

Training dataset: 10000 vectors Testing dataset: 10000 vectors
Tensorflow and Keras frameworks were used
The neural network model has:

- input layer with 38 neurons
- one hidden layer with 20 neurons and relu activation function
- output layer with 1 neuron and softmax activation function



Thresholds for unacceptable risk

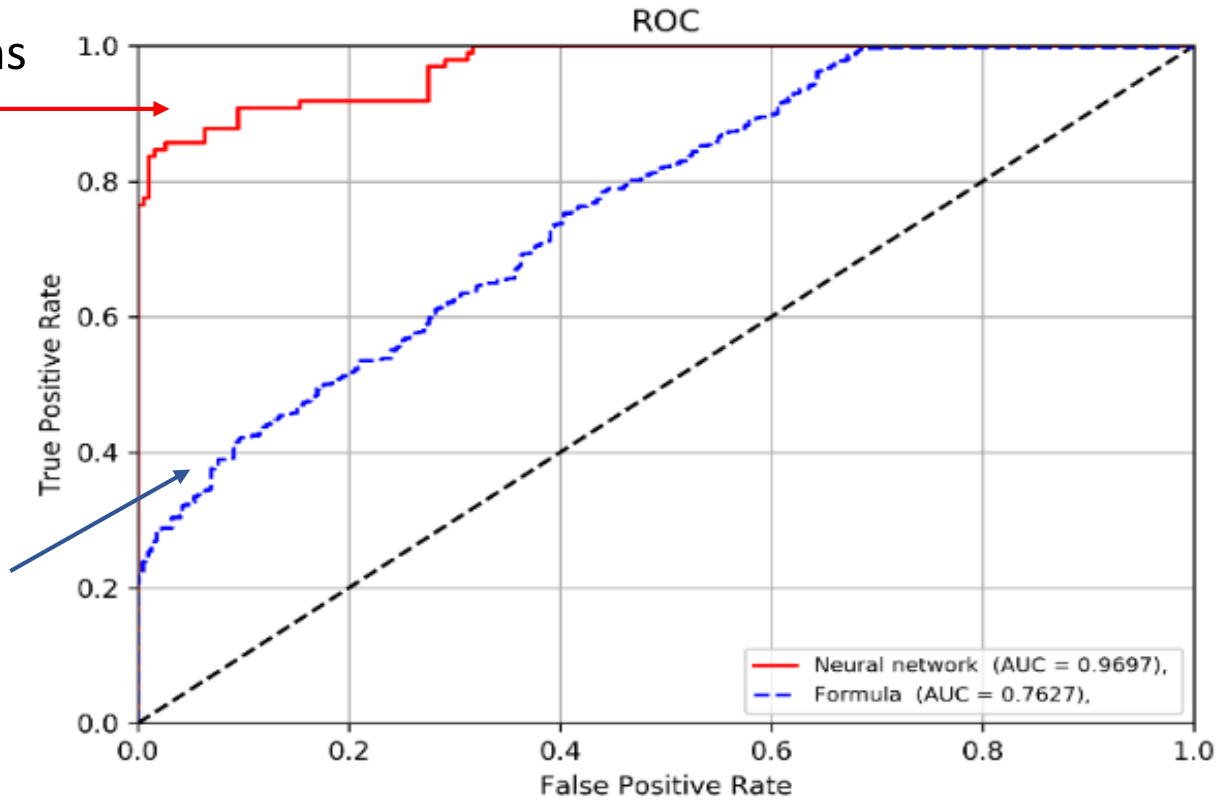
Asset type	Network type	Permissible probability of node failure
Smart phone	MANET	< 1%
Laptop	MANET	< 0.5%
Vehicle	VANET	< 0.01%
Traffic light	VANET	< 0.1%
Road-side unit	VANET	< 1%
Smart door lock	IoT	< 3%
Medical sensor	IoT	< 0.03%
Temperature sensor	IIoT	< 0.01%
Database server	IIoT	< 0.1%
Smart robot	IIoT	< 0.1%

Neural network input parameters

Parameter	Description
Device number	Device ID (0-10000)
Device type	Device type. For example: mobile, vehicle, traffic light, smart door lock, medical sensor, … (0-10)
Q_{Tk}	Device cost in US dollars (100-50000)
Associated with T_i (n times)	The device is associated with with T_i (0 or 1)
Probability of BH	Probability of Black Hole attack (0-100)
Probability of GH	Probability of Gray Hole attack (0-100)
Probability of DoS	Probability of DoS attack (0-100)
Probability of DDoS	Probability of DDoS attack (0-100)
Probability of WH	Probability of Wormhole attack (0-100)
I_{TiTk} (n times)	Coefficient of influence of devices of the type $T_i \in T$ on devices of the type $T_k \in T$ (0-1)
C_{TiTk} (n times)	Coefficient showing the number of devices of type $T_k \in T$ with which device of type $T_i \in T$ interacts (0-N)

Experimental results

Maximum classification accuracy was achieved with the following neural network parameters: 3 layers, 40 epochs of training, a training set equal to 10000, and is 97%.



$$R(U_j)T_i = P(U_j)\sum_{k=1}^{n} I_{T_iT_k} C_{T_iT_k} \times Q_{T_k}$$

where $R(U_j)T_i$ is a security risk when implementing the threat $U_j \in U$ for the device of the type $T_i \in T$; $P(U_j)$ is a probability of realization of the threat $U_j \in U$; $I_{T_iT_k}$ is a coefficient of influence of devices on each other; $C_{T_iT_k}$ is a coefficient of the number of the device interactions with each other; Q_{T_k} is an amount of possible damage.

Conclusion

1. An analysis of existing works showed that researchers are actively looking for new approaches to risk assessment, since traditional methods are not able to work in a rapidly changing environment.
2. A neural network model was developed, namely a three-layer perceptron, which was trained on labeled data, and then the classification quality was assessed on unlabeled data.. All assets were typed, and a threshold of the acceptable level of risk was determined for each asset type.
3. The test results showed an accuracy of 97%, which speaks of the promise of the proposed approach.
4. The main advantages of the proposed approach are: the ability to work in rapidly changing conditions, high classification accuracy when working with big data, the possibility of dynamic risk assessment, as well as the ability to work in conditions of limited awareness of the state of the entire smart city network.
5. In the future, it is planned to continue research in this area:

- add new features to datasets (for example, various network indicators: the ratio of sent and lost packets, throughput, number of hops, etc., as well as economic indicators: ROI, ROA, ROE);
- rank the risk (not binary: acceptable and unacceptable, but multi-level: high-level, high, moderate, average, low);
- compare proposed neural network approach with other existing cybersecurity risk assessment methods.