Technique for Finding and Investigating the Strongest Combinations of Cyberattacks on Smart Grid Infrastructure

Igor Kotsiuba
G.E. Pukhov Institute for Modeling in Energy Engineering, National Academy of Sciences of Ukraine
Kyiv, Ukraine
i.kotsiuba@gmail.com

Inna Skarga-Bandurova
School of Engineering, Computing and Mathematics
Oxford Brookes University
Oxford, United Kingdom
iskarga-bandurova@brookes.ac.uk
ORCID ID: 0000-0003-3458-8730

Alkiviadis Giannakoulias
School of Electrical and Computing Engineering
National Technical University of Athens
Athens, Greece
alkiviadis.giannakoulias@eurodyn.com

Mykhailo Chaikin
CRDF Global
Kyiv, Ukraine
mchaikin@crdfglobal.org

Aleksandar Jevremovic
IEEE member
Serbia
ajevremovic@ieee.org

***Abstract*— Recently, smart grids have become a vector of the energy policy of many countries. Due to structural and operation features, smart grids are a constant target of combined and simultaneous cyberattacks. To maximize security and to optimize existing network schemes to prevent cyber intrusion, in this paper, we propose an approach to decision support in finding and identifying the most potent attack combinations that can set the system to maximum damage. The main purpose is to identify the most severe combinations of attacks on smart grid components that potentially can be implemented from the perspective of the attacker. In this context, the problem of finding weaknesses points in the network configuration of a smart grid and assessing the impact of events on cyberinfrastructure is considered. The technique for detecting and investigating the strongest combinations of cyberattacks on the smart grid network is given with an example of the analysis of the spread of pandemic software in a system with arbitrary structure.**

Keywords*—*smart grid, cyber-physical system, cyberattack, network, graph, malware, virus

I. INTRODUCTION

Cyber risk is a unique problem for the smart grid infrastructure since a cyberattack can easily move from the cyber sphere to the physical world. Talking about the cyberattacks in distributed intelligent electrical power systems, the following notes should be taken into account:

- Several types of attacks can be launched simultaneously in the cyber-physical system of the smart grid infrastructure.
- Cybercriminals create various attacks depending on the simplicity of actions, the course of events, and less complexity in creating an attack to maximize harm.
- Given this, existing cyber threats should be considered both in the plane of physical components and in terms of related ICT components.

The most representative recent examples of the implementation of cyber threats are the blackouts in Ukraine in 2015 and 2016. They occurred as a result of a series of cyberattacks, as well as successful attacks by a group of hackers Dragonfly 2.0, which in 2017 gained access to several network interfaces of energy companies that are leveraged by operators for transmitting commands to equipment, such as circuit breakers.

A list of the most prominent known cybersecurity incidents in the last 15 years is given in Table I.

TABLE I. KNOWN CYBERSECURITY INCIDENTS AGAINST ENERGY GRID INFRASTRUCTURE

Year	Name of the object or malware	Type/Method	Target/ Impact
2003	NPP Davis-Bessie	Worm	Network Disruption
2003	Venezuela Maritime Terminal	Virus	Disruption of service
2007	Tehanama Kolusa Channel Management	Compromise threat	Unknown
2008	Attack on Power Supply of four US Cities	Unknown	Service Disruption, Money Demand
2009	US Electrical Network	Infrastructure Mapping	Unknown
2010	Stuxnet	Worm, Trojan	Unauthorized access to SCADA, Service disruption, Destruction of equipment
2011	Night Dragon	Social engineering, user compromise, Phishing, Windows exploits	Unauthorized access to the control system and information system
2011	Duqu	Virus, Windows-exploits	Data theft, Information assets of industrial control systems
2012	Greek oil company	Trojan, Social Engineering	Cyber-espionage
2012	Aramco oil company	Virus	Cyber-espionage
2012	Flame	Worm, Windows-exploits	Cyber-espionage
2013	US energy companies in the Middle East	Multiple attacks, details not available	Sabotage
2013	Austrian and German power grids	Unexpected DDoS	Operations were not affected
2014	Dragonfly	Worm, Windows-exploits, Trojan, Backdoors	Cyber-espionage
2015-2016	Ukrainian electricity networks	Compromise threat Trojan, Worm	Service disruption
2017	Dragonfly 2.0	Phishing,	Data collection and

I. Kotsiuba, I. Skarga-Bandurova, A. Giannakoulias, M. Chaikin, and A. Jevremovic, "Technique for Finding and Investigating the Strongest Combinations of Cyberattacks on Smart Grid Infrastructure," in 2019 IEEE International Conference on Big Data (Big Data), 2019, doi: 10.1109/bigdata47090.2019.9006335

The published version is available at IEEE Xplore Digital Library: https://ieeexplore.ieee.org/document/9006335

		malicious email attachments, Trojan	exploration, unauthorized access
2017	Maersk. Companies and port terminals around the world	NotPetya, Windows exploits	Service disruption.
2018	DisTrack	Virus	Service disruption

In 2017, the distribution of the NotPetya ransomware occurred at a transnational level. This intervention was recognised as the act of a full-scale cyber-war. A few hours after its first appearance, the worm went beyond Ukraine and intervened in many machines around the world "from a hospital in Pennsylvania to the Tasmanian chocolate factory" [1]. Recent attacks on Maersk (2017) and DisTack (2018) lead to immense losses disrupting services in company networks all around the world. In April 2018, as a result of a cyberattack, four U.S. pipeline companies experienced a shutdown of their electronic systems that lasted for several days.

The list of ICT components of smart energy networks that should be considered as a potential source of vulnerabilities should include [2]:

- Operating systems and components: generators, transformers, supervisory control and data acquisition systems (SCADA), energy management / distribution systems (EMS / DMS), programmable logic controllers (PLCs), substations, smart meters and other smart electrical devices.
- Classic IT systems: PCs, servers, mainframes, applications, databases, websites, web services, etc.
- Networks and communication protocols: Ethernet, Wi-Fi, PRIME, DLMS / COSEM, Zigbee, 4G, DNP3, etc.
- Endpoints: smart meters, EV, smartphones and other mobile devices.

II. MODELS OF MALWARE

Table II presents the types of cyber-physical attacks on smart grid components in terms of their impact on integrity, privacy, and accessibility.

TABLE II. CLASSIFICATION OF CYBER-PHYSICAL ATTACKS

Attack model	Cyber	Physical	The target feature
Denial of Service (DoS)	+	+	Availability
Listening	+		Confidentiality
False Data Injection (FDI)	+	+	Integrity
Insert malware	+	+	Authentication
Людина в середині	+	+	Integrity, Confidentiality
Enemy device		+	Integrity, Confidentiality
Unauthorized access		+	Authorization, Confidentiality
Wireless scrambling (encryption, capture)	+		Integrity

According to the classification given in [3], the characteristics and parameters of malware could be represented in terms of the pandemic, endemic, and infectious types. (see Table III). Contrary to [3], where the characteristics of malware distribution in networks are studied based on mathematical modeling, this study defines the components used to model the initial stages of malware distribution in such networks to analyze the availability of network branches to cyberattacks.

TABLE III. MODELS OF MALWARE

Characteristics	Types of malware		
	Pandemic	Endemic	Infectious
Scan type	Aggressive topological	Hit-list	Passive
Distribution	Infected Node Scans the Whole Subnet	Infected Node Scans Part of the Subnet	Infected Node infects the network through an established communication channel
Scanning velocity, sec	100	1	-
Load, byte	500	5000	5000
Morphism	Oligomorphic	Polymorphic	Metamorphic
Complexity	Simple	Complex	Complex
Distribution	Self-spreading	Self-spreading	Built-in
Examples	Red 1, 2 [4,5], Nimda [4], Slammer [6, 7], Conficker [8,9]	Regin [10], Duqu [11,12], Flame [11].	Gauss [11], Equation [13], AdWind [14,15], Grey Energy [16]

It is worth noting that the taxonomy presented in Table III is conditional since malware can combine different types of characteristics, for example, Stuxnet virus [17] has many functions similar to the endemic category, but it does not use hit lists. The speed and scale of distribution of WannaCry [18] and NotPetya [19], which in 2017 infected more than 200,000 devices in more than 100 countries, causing more than $ 4 billion in losses in just 24 hours [20], makes it pandemic in while they use more sophisticated implementation and anti-tampering methods, such as XOR encryption and fake Microsoft digital signatures [21]. The same statement applies to their predecessor, Petya, whose developers used the methods commonly used by penetration testers and hackers and built sophisticated multi-threaded automation of these methods in one piece of code.

The malware of GreyEnergy currently has no destructive capabilities and seems to focus on spyware and intelligence operations on control system workstations that work with SCADA software and servers, giving it a reason to classify it as a third category. However, GreyEnergy has a modular architecture, which means that its capabilities can be expanded [22]. Also, ESET experts note that GreyEnergy has been involved in attacks on energy companies and other Ukrainian and Polish value units over the last three years.

Identification of potential high-risk vulnerabilities is a vital component of an advanced security strategy and should be part of the overall smart grid infrastructure management program. Early warning systems, risk management analytics, security monitoring systems and digital audit systems enable businesses and researchers to make better decisions. Traditional decision support techniques rely heavily on data analytics algorithms for security, cyberattack capabilities and vectors, data breaches, and more.

However, the regulatory analysis showed that due to the technical difference between the IEs and the CVSS used by NIST, threat assessment could only be performed on IEC61850 computer nodes such as database servers, engineering stations, human-machine interfaces and

gateways. Therefore, it is necessary to develop and use a new metric scheme capable of taking into account different levels of threats and auditing the security of smart energy systems under metric schemes.

In view of the above, in this work, we assumed that several simultaneous attacks on the similar smart grid system components could occur. The main purpose is to identify the strongest combinations of attacks on smart grid components that potentially can be implemented from the perspective of the attacker.

III. BASIC ASSUMPTIONS

One of the most challenging issues caused by cyber intrusions is cascading network outages due to the simultaneous attack on several nodes of a distributed power system. In this case, the goal of a cyberattack is to disconnect or switch network branches from an operational state to shutdown. Potential attacks on the network ci form the combination vector C: C = {c_1, c_2, c_3, c_4 ..., c_m}. The elements of this vector include the sets of sequences of simultaneous disconnection of certain network branches and enable calculating the damage caused to the system due to the implementation appropriate combinations of cyberattacks. Then the time to turn off the power system t(ci) forms another vector: T = {t (c_1), t (c_2), t (c_3), t (c_4), ..., t (c_m)}.

Using the introduced notations, a combinational attack is considered as powerful if a maximum number of disconnections of the branches of the system k → E can be achieved in a minimum time t (c_i) → min.

For further consideration, the following assumptions are made:

- Assumption 1: Any type of attack can be used to attack simultaneously.
- Assumption 2: All an attacker needs is to manipulate a relay or switch. A line switching attack can, for example, be triggered by initiating an emergency protection scheme or a corrective action scheme.
- Assumption 3: Attackers have the resources to attack multiple lines at the same time to initiate a simultaneous attack.
- Assumption 4: For n-k unforeseen circumstances, there should be a kmax value of N. It is assumed that the maximum kmax branches will fire simultaneously with a simultaneous attack.
- Assumption 3: The largest attack combination is considered for the minimum time required to achieve a power outage, calculated as a percentage of power lost (MW).
- Assumption 4: The maximum blackout is 100% power loss.

In this case, the percentage of lost MW is calculated using the formula below:

$$\%\ \text{MW lost} = \frac{a-b}{a} \cdot 100\%,$$

where *a* is the total number of MW before the attack and *b* is the total number of MW after the attack.

The problem of finding weaknesses in the configuration of a distributed electricity network is formulated in the form of a connection prediction problem in its cyber graph.

For a non-directional graph G (V, E), where V is the set of nodes i, E is the set of edges e (i, j) connecting the nodes i and j, it is necessary to determine the distance function between nodes of the graph, which will guarantee for the structure of the graph G (t_0, t_0*), given in the interval of time (t_0, t_0*), to predict the structure of G (t_1, t_1*) in the interval (t_1, t_1*).

Due to the features of smart grid involve a dynamic network structure the PageRank importance indicator, which is commonly used for various social network prediction tasks [23], on the Internet, and cyber threat detection through link analysis [24] can be used for assessing the impact of events on cyber infrastructure of distributed power grids. The reasons for using the PageRank algorithm for calculating node criticality are:

- The results are calculated using a stochastic approach that reflects the randomness in the evolution of the model. With regard to cyber-threat infrastructures, we assume that there is a constant evolution of smart grid cyberspace, in the form of new organizations, owners, IPs, servers, malware samples, domains, and registrars. This emergence of new peaks influences the evolution of network accessibility estimates for cybercriminals.
- The random model illustrates the access to the nodes of the graph with probability (damping factor).

Similar to changing cyber-threats infrastructure, the use of a probabilistic approach model is interesting because it enables to track potential actions taken through infected machines. For example, it is assumed that the compromised domain can be visited by the infected machine, the IP address can be connected to infected machines or connected to the server of the compromised domain, the FTP server can be used to download the stolen information, the SMTP server can be used to start spam or phishing campaigns, you can use the IRC channel to instruct bots to launch DDoS attacks, distribute malware, or other malicious activity, and more.

IV. TECHNIQUE FOR ASSESSING THE IMPACT OF EVENTS ON SMART GRID CYBER INFRASTRUCTURE

Given the scale and constantly changing the structure of smart grid networks, it is additionally suggested that attackers have the ability to attack any line by infecting the target node. Therefore, it is believed that both the purpose of the attack and the purpose of protection is to find the peaks that have the greatest impact on the infrastructure. Decisions about the strongest combinations of attacks capable of giving the system maximum damage are made by calculating the importance of the nodes of the graph and redistributing them according to the values obtained.

The analysis can be based on cyber graph models or a topological network diagram, which is also extrapolated as a graph.

Step 1. Create a matrix of branches.

The adjacency matrix of graph G with a finite number of nodes n is a square matrix of size n × n, which is formed from the elements of nodes whose values are equal to the weight ω_{ij} of the edge e (i, j).

$$M = \begin{pmatrix} \omega_{11} & \cdots & \omega_{1n} \\ \vdots & \ddots & \vdots \\ \omega_{n1} & \cdots & \omega_{nn} \end{pmatrix}.$$

Step 2. Generate a combinational vector of attacks.

For k disconnections with E edges, the number of possible combinations is calculated by the following formula [25]:

$$c = \binom{E}{k} = \frac{E!}{k!(E-k)!}.$$

Step 3. Create a combination matrix of transitions.

A stochastic matrix is created for each attack vector, in which all columns are rows of real numbers from 0 to 1, giving in the sum 1:

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{pmatrix},$$

where the value of each element of the matrix is determined by:

$$w_{ij} = \begin{cases} \dfrac{1}{d_0(v_j)}, & \forall i, j \quad i \in S(j), \\ 0 & \forall i, j \quad i \notin S(j). \end{cases}$$

Step 4. Calculate the importance of the nodes

The importance of PR (v_i) of the node associated with node v_i and $d_o(v_i)$ - the number of output edges from node v_i is calculated as [26]:

$$PR(v_i) = d\left[\sum_{v_i \in I(v_i)} \frac{PR(v_i)}{\mathrm{d}_o(v_i)}\right] + (1-d)\frac{1}{|D|},$$

where d is the damping factor (0.85).

In this case, it is assumed that if the attacker has reached node v_i with probability d, then the probability of reaching another node v_i is $1/d_o(v_j)$.

As a result, one equation for a node with the corresponding number of unknown values of PR (v_i) can be obtained in one step.

Assuming that $\sum_{i=1}^{n} PR(v_i) = 1$, the algorithm iteratively finds different values for each t until the values converge, i.e.

$$| (PR (vi), t) - (PR (vi), t – 1) | 1 <\varepsilon,$$

where ε is the permissible error.

Step 5. Redistribute the nodes of the graph by their criticality for the invasion.

At this stage, the nodes of the graph are sorted by PR and decisions are made on what to do next.

The resulting model reflects access to elements of cyberinfrastructure and can be used to analyze potential attacks through infected channels.

Step 6. Analyze the impact of cyberattack combinations on the smart grid infrastructure.

The analysis calculates the possible damage and identifies the strongest combinations of attacks.

V. CASE STUDY

The proposed approach was tested for analysis of the distribution of pandemic software in the smart grid network with arbitrary structure. To evaluate the possibilities and potential of the proposed method, we consider a hypothetical network diagram (Fig. 1).

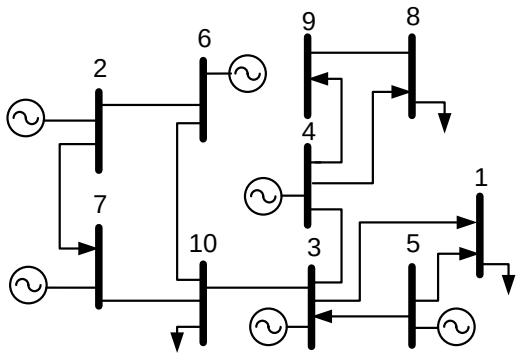

Figure 1 – An example of a network diagram

The network structure has 10 buses, 6 transformers and 14 transmission lines, which are given unique numbers corresponding to the connection points of the respective i-j buses. Taking into account accepted assumptions, this scheme can be represented in the form of a graph (Fig. 2) for which the corresponding adjacency and incident matrix is formed.

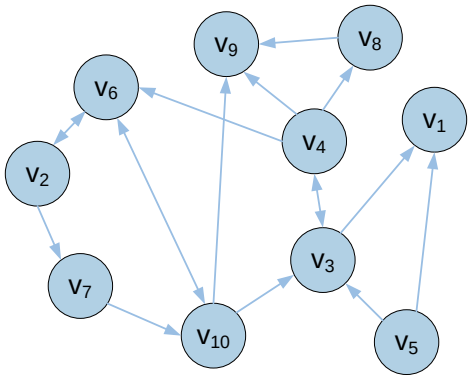

Figure 2 - Graph of the system under study

The adjacency matrix of the graph

v_i	1	2	3	4	5	6	7	8	9	10
1	0	0	1	0	1	0	0	0	0	0

2	0	0	0	0	0	1	1	0	0	0
3	1	0	0	1	1	0	0	0	0	1
4	0	0	1	0	0	1	0	1	1	0
5	1	0	1	0	0	0	0	0	0	0
6	0	1	0	1	0	0	0	0	0	1
7	0	1	0	0	0	0	0	0	0	1
8	0	0	0	1	0	0	0	0	1	0
9	0	0	0	0	0	0	0	1	0	1
10	0	0	1	0	0	1	1	0	1	0

Input incidence matrix

v_i	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	1	0	0	0
3	1	0	0	1	0	0	0	0	0	0
4	0	0	1	0	0	1	0	1	1	0
5	1	0	1	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	1
7	0	0	0	0	0	0	0	0	0	1
8	0	0	0	0	0	0	0	0	1	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	1	0	0	1	0	0	1	0

Then, for the graph shown in Fig. 2, the sets of reach of each node can be represented as follows:

$$R(v_i) = \{v_i\} \cup \Upsilon^{+1}(v_i) \cup \Upsilon^{+2}(v_i) \cup \ldots \cup \Upsilon^{+p}(v_i).$$

$$R^{Out}(v_1) = \{v_1\}.$$

$$R^{Out}(v_2) = \{v_2\} \cup \{v_2, v_6\} \cup \{v_2, v_7\} \cup \{v_2, v_7, v_{10}\} \cup \{v_2, v_7, v_{10}, v_9\} \cup \{v_2, v_7, v_{10}, v_3\} \cup \cup \{v_2, v_7, v_{10}, v_3, v_4\} \cup \{v_2, v_7, v_{10}, v_3, v_4, v_6\} \cup \{v_2, v_7, v_{10}, v_3, v_4, v_9\} \cup \cup \{v_2, v_7, v_{10}, v_3, v_4, v_8\} \cup \{v_2, v_7, v_{10}, v_3, v_1\} = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

$$R^{Out}(v_3) = \{v_3\} \cup \{v_3, v_1\} \cup \{v_3, v_4\} \cup \{v_3, v_4, v_3, v_1\} \cup \{v_3, v_4, v_6\} \cup \{v_3, v_4, v_6, v_2\} \cup \cup \{v_3, v_4, v_6, v_2, v_7\} \cup \{v_3, v_4, v_6, v_2, v_7, v_{10}\} \cup \{v_3, v_4, v_6, v_2, v_7, v_{10}, v_9\} \cup \cup \{v_3, v_4, v_6, v_2, v_7, v_{10}, v_3\} \cup \{v_3, v_4, v_6, v_2, v_7, v_{10}, v_3, v_1\} \cup \{v_3, v_4, v_8\} \cup \{v_3, v_4, v_9\} = = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

$$R^{Out}(v_4) = \{v_4\} \cup \{v_4, v_3\} \cup \{v_4, v_3, v_1\} \cup \{v_4, v_6\} \cup \{v_4, v_6, v_2\} \cup \{v_4, v_6, v_2, v_7\} \cup \cup \{v_4, v_6, v_2, v_7, v_{10}\} \cup \{v_4, v_6, v_2, v_7, v_{10}, v_9\} \cup \{v_4, v_6, v_2, v_7, v_{10}, v_3\} \cup \cup \{v_4, v_6, v_2, v_7, v_{10}, v_3, v_1\} \cup \{v_4, v_8\} \cup \{v_4, v_9\} = = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

$$R^{Out}(v_5) = \{v_5\} \cup \{v_5, v_1\} \cup \{v_5, v_3\} \cup \{v_5, v_3, v_4\} \cup \{v_5, v_3, v_4, v_6\} \cup \{v_5, v_3, v_4, v_8\} \cup \cup \{v_3, v_4, v_9\} = \{v_1, v_3, v_4, v_5, v_6, v_8, v_9\}.$$

$$R(v_6) = \{v_6\} \cup \{v_6, v_2\} \cup \{v_6, v_{10}\} \cup \{v_6, v_2, v_7\} \cup \{v_6, v_2, v_7, v_{10}\} \cup \{v_6, v_2, v_7, v_{10}, v_9\} \cup \cup \{v_6, v_2, v_7, v_{10}, v_3\} \cup \{v_6, v_2, v_7, v_{10}, v_3, v_4\} \cup \{v_6, v_2, v_7, v_{10}, v_3, v_4, v_9\} \cup \cup \{v_6, v_2, v_7, v_{10}, v_3, v_4, v_8\} \cup \{v_6, v_2, v_7, v_{10}, v_3, v_4, v_8, v_9\} \cup \{v_6, v_2, v_7, v_{10}, v_3, v_1\} = = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

$$R^{Out}(v_7) = \{v_7\} \cup \{v_7, v_{10}\} \cup \{v_7, v_{10}, v_6\} \cup \{v_7, v_{10}, v_6, v_2\} \cup \{v_7, v_{10}, v_9\} \cup \{v_7, v_{10}, v_3\} \cup \cup \{v_7, v_{10}, v_3, v_4\} \cup \{v_7, v_{10}, v_3, v_4, v_9\} \cup \{v_7, v_{10}, v_3, v_4, v_8\} \cup \{v_7, v_{10}, v_3, v_4, v_8, v_9\} \cup \cup \{v_7, v_{10}, v_3, v_1\} = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

$$R^{Out}(v_8) = \{v_8\} \cup \{v_8, v_9\}.$$

$$R^{Out}(v_9) = \{v_9\}.$$

$$R^{Out}(v_{10}) = \{v_{10}\} \cup \{v_{10}, v_6\} \cup \{v_{10}, v_9\} \cup \{v_{10}, v_3\} \cup \{v_{10}, v_6, v_2\} \cup \{v_{10}, v_6, v_2, v_7\} \cup \cup \{v_{10}, v_3, v_1\} \cup \{v_{10}, v_3, v_4\} \cup \{v_{10}, v_3, v_4, v_9\} \cup \{v_{10}, v_3, v_4, v_8\} \cup \{v_{10}, v_3, v_4, v_8, v_9\} \cup = = \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{10}\}.$$

The output incidence matrix

v_i	1	2	3	4	5	6	7	8	9	10
1	0	0	1	0	1	0	0	0	0	0
2	0	0	0	0	0	1	0	0	0	0
3	0	0	0	0	1	0	0	0	0	1
4	0	0	0	0	0	1	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	1	0	0	0	0	0	1
7	0	1	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0	0	0
9	0	0	0	1	0	0	0	1	0	0
10	0	0	0	0	0	1	1	0	0	0

Sets for input streams:

$$R^{In}(v_1) = \{v_1\} \cup \{v_3, v_1\} \cup \{v_5, v_1\} \cup \{v_4, v_3, v_1\} \cup \{v_{10}, v_3, v_1\} \cup \{v_6, v_{10}, v_3, v_1\} \cup \cup \{v_2, v_6, v_{10}, v_3, v_1\} \cup \{v_7, v_{10}, v_3, v_1\} \cup \{v_2, v_7, v_{10}, v_3, v_1\} = = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_{10}\}.$$

$$R^{In}(x_2) = \{x_2\} \cup \{x_6, x_2\} \cup \{x_4, x_6, x_2\} \cup \{x_{10}, x_6, x_2\} \cup \{x_3, x_4, x_6, x_2\} \cup \cup \{x_{10}, x_3, x_4, x_6, x_2\} \cup \{x_5, x_3, x_4, x_6, x_2\} = \{x_2, x_3, x_4, x_5, x_6, x_{10}\}.$$

$$R^{In}(v_3) = \{v_3\} \cup \{v_4, v_3\} \cup \{v_5, v_3\} \cup \{v_{10}, v_3\} \cup \{v_6, v_{10}, v_3\} \cup \cup \{v_2, v_6, v_{10}, v_3\} \cup \{v_4, v_6, v_{10}, v_3\} = \{v_2, v_3, v_4, v_5, v_6, v_{10}\}.$$

$$R^{In}(v_4) = \{v_4\} \cup \{v_3, v_4\} \cup \{v_5, v_3, v_4\} \cup \{v_{10}, v_3, v_4\} \cup \{v_6, v_{10}, v_3, v_4\} \cup \cup \{v_2, v_6, v_{10}, v_3, v_4\} \cup \{v_2, v_7, v_3, v_4\} = \{v_2, v_3, v_4, v_5, v_6, v_7, v_{10}\}.$$

$$R^{In}(v_5) = \{v_5\} \cup \{v_5, v_1\} \cup \{v_5, v_3\} \cup \{v_5, v_3, v_4\} \cup \{v_5, v_3, v_4, v_6\} \cup \{v_5, v_3, v_4, v_8\} \cup \cup \{v_3, v_4, v_9\} = \{v_1, v_3, v_4, v_5, v_6, v_8, v_9\}.$$

$$R^{In}(v_6) = \{v_6\} \cup \{v_4, v_6\} \cup \{v_{10}, v_6\} \cup \{v_2, v_6\} \cup \{v_7, v_{10}, v_6\} \cup \{v_2, v_7, v_{10}, v_6\} \cup \cup \{v_3, v_4, v_6\} \cup \{v_{10}, v_3, v_4, v_6\} \cup \{v_7, v_{10}, v_3, v_4, v_6\} \cup \{v_2, v_7, v_{10}, v_3, v_4, v_6\} \cup \cup \{v_5, v_3, v_4, v_6\} = \{v_2, v_3, v_4, v_5, v_6, v_7, v_{10}\}.$$

$$R^{In}(v_7) = \{v_7\} \cup \{v_2, v_7\} \cup \{v_6, v_2, v_7\} \cup \{v_{10}, v_6, v_2, v_7\} \cup \{v_4, v_6, v_2, v_7\} \cup \cup \{v_3, v_4, v_6, v_2, v_7\} \cup \{v_{10}, v_3, v_4, v_6, v_2, v_7\} \cup \{v_5, v_3, v_4, v_6, v_2, v_7\} = = \{v_2, v_3, v_4, v_5, v_6, v_7, v_{10}\}.$$

$$R^{In}(v_8) = \{v_8\} \cup \{v_4, v_8\} \cup \{v_3, v_4, v_8\} \cup \{v_5, v_3, v_4, v_8\} \cup \{v_{10}, v_3, v_4, v_8\} \cup \cup \{v_6, v_{10}, v_3, v_4, v_8\} \cup \{v_2, v_6, v_{10}, v_3, v_4, v_8\} \cup \{v_2, v_7, v_3, v_4, v_8\} = = \{v_2, v_3, v_4, v_5, v_6, v_7, v_8, v_{10}\}.$$

$$R^{In}(v_9) = \{v_9\} \cup \{v_4, v_8\} \cup \{v_{10}, v_9\} \cup \{v_3, v_4, v_9\} \cup \{v_5, v_3, v_4, v_9\} \cup \{v_{10}, v_3, v_4, v_9\} \cup \cup \{v_6, v_{10}, v_3, v_4, v_9\} \cup \{v_2, v_6, v_{10}, v_3, v_4, v_9\} \cup \{v_2, v_7, v_3, v_4, v_9\} = = \{v_2, v_3, v_4, v_5, v_6, v_7, v_9, v_{10}\}.$$

$R^{In}(v_{10}) = \{v_{10}\} \cup \{v_7, v_{10}\} \cup \{v_6, v_{10}\} \cup \{v_2, v_7, v_{10}\} \cup \{v_6, v_2, v_7, v_{10}\} \cup$
$\cup \{v_4, v_6, v_{10}\} \cup \{v_3, v_4, v_6, v_{10}\} \cup \{v_3, v_4, v_6, v_2, v_7, v_{10}\} \cup \{v_5, v_3, v_4, v_6, v_2, v_7, v_{10}\} =$
$= \{v_2, v_3, v_4, v_5, v_6, v_7, v_{10}\}.$

As can be seen from the above, the sets of reachability alone do not make it possible to decide on the importance of the peaks in terms of their availability for the pandemic distribution of malicious software.

For the scheme in Fig. 1, the second step gives us 24 combinations of attacks for a single shutdown. Considering the unpredictability of E-2 (assuming k = 2), the number of possible combinations of simultaneous attacks will be [27]:

$$c(E-2) = \frac{E!}{2!(E-2)!} = \frac{E(E-1)}{2} = \frac{E^2 - E}{2} = \frac{14^2 - 14}{2} = 189,$$

which actually proves how much damage can be done to power systems using well-planned attacks.

The next step of the proposed method is to create combinatorial matrices. The formation of combinational transition matrices is performed according to the incident matrices obtained in the previous step. As a result, we get the following distributions of conversion values.

Input transition matrix

v_i	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1/3	1	0	0	0
3	0,5	0	0	1	0	0	0	0	0	0
4	0	0	1/3	0	0	1/3	0	1	1/3	0
5	0,5	0	1/3	0	0	0	0	0	0	0
6	0	1	0	0	0	0	0	0	0	0,5
7	0	0	0	0	0	0	0	0	0	0,5
8	0	0	0	0	0	0	0	0	1/3	0
9	0	0	0	0	0	0	0	0	0	0
10	0	0	1/3	0	0	1/3	0	0	1/3	0

Output transition matrix

v_i	1	2	3	4	5	6	7	8	9	10
1	0	0	1	0	0,5	0	0	0	0	0
2	0	0	0	0	0	1/3	0	0	0	0
3	0	0	0	0	0,5	0	0	0	0	0,5
4	0	0	0	0	0	1/3	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0
6	0	0	0	1/3	0	0	0	0	0	0,5
7	0	1	0	0	0	0	0	0	0	0
8	0	0	0	1/3	0	0	0	0	0	0
9	0	0	0	1/3	0	0	0	1	0	0
10	0	0	0	0	0	1/3	1	0	0	0

The next step is computing the importance of the nodes, which is determined by the number of output edges, taking into account their connectivity. This task is iterative and causes the use of granulation to achieve maximum processing speed with the maximum degree of systematic abstraction [28].

According to [29], the granular processing principle involves solving a problem for a single node and then extrapolating the output to the nodes of the entire graph. The algorithm of calculation thus coincides with six stages [30].

Stage 1. Select a random node v_i.

Stage 2. Compute all edges included in this node:

$$PR(v_i) = \sum_{j=1}^{k} e_{ij},$$

where $e_{i,j} = 1$ if node v_i is the end of an arc joining nodes v_i and v_j, and $e_{i,j} = 0$ otherwise.

Stage 3. For each source of input edges, the total number of output edges $d_0(v_j)$ is calculated:

$$d_0(v_j) = \sum_{j=1}^{k} e_{ij},$$

where $e_{i,i} = 1$, if node v_i is the beginning of an arc connecting nodes v_i and v_j, and $e_{i,j} = 0$ otherwise.

Stage 4. Compute the importance of the PR (v_i) of the node associated with the node v_i, where $d_o(v_j)$ is the number of output edges from the node v_j.

Stage 5. For all nodes repeat Stages 1-3.

Stage 6. Check the convergence of results. The calculations are completed when the conditions are reached $\sum_{i=1}^{n} PR(v_i) = 1$, otherwise stages 1-5 are repeated. It is believed that convergence occurs when all ranks are within the error boundary $\left|PR(v_i)_{t+1} - PR(v_i)_t\right| < \varepsilon,$ where ε is the error limit, an arbitrary value from 0 to 1. The smaller the error limit, the more accurate the result.

TABLE IV. THE RESULT OF CALCULATING THE IMPORTANCE OF NODES BY PR

v_i	In-Deg	Out-Deg	Deg	Weighted In-Deg	Weighted Out-Deg	Weighted Degree	$PR(v_i)$
1	2	0	2	2	0	2	0.077145
2	1	2	3	1	2	3	0.077640
3	3	2	5	3	2	5	0.138118
4	1	4	5	1	4	5	0.134641
5	0	2	2	0	2	2	0.077145
6	3	2	5	3	2	5	0.105522
7	1	1	2	1	1	2	0.076924
8	1	1	2	1	1	2	0.072951
9	3	0	3	3	0	3	0.103558
10	2	3	5	2	3	5	0.136356

After ordering the vertices by PR, we obtain the following sequence AR (see Table V):

$$v_3 \succ v_{10} \succ v_4 \succ v_6 \succ v_9 \succ v_2 \succ v_1 \succ v_5 \succ v_7 \succ v_8$$

TABLE V. RESULT OF THE REDISTRIBUTION OF NODES BY THEIR IMPORTANCE

AR	v_i	In-Deg	Out-Deg	Deg	Weight In-Deg	Weight Out-Deg	Weight Degree	PageRank
1	3	3	2	5	3	2	5	0.138118
2	10	2	3	5	2	3	5	0.136356
3	4	1	4	5	1	4	5	0.134641
4	6	3	2	5	3	2	5	0.105522

5	9	3	0	3	3	0	3	0.103558
6	2	1	2	3	1	2	3	0.07764
7	1	2	0	2	2	0	2	0.077145
8	5	0	2	2	0	2	2	0.077145
9	7	1	1	2	1	1	2	0.076924
10	8	1	1	2	1	1	2	0.072951

VI. DISCUSSION

As a result of the redistribution, it becomes possible to determine the order of verification of nodes by their importance / criticality for the system as a whole. Another option to apply the proposed approach and its natural evolution is to evaluate the strongest combinations of network attacks. In this context, it is possible to utilise relatively recently presented approaches [25, 31], which enable the modeling of attacks on individual branches of the graph and assessing the potential harm from their implementation.

As a rule, the electrical topological structure of the energy system is static (updated only when new elements are included) and contains detailed information about system assets, their characteristics and configurations. The only dynamically changing parameter is the state of the network switching equipment. Obviously, in a power system, the status of network switches only changes when the system is reconfigured, which also does not happen often, while the modules of the power management system perform their operations very often, even every few minutes, depending on the program. With this in mind, the best target for cyberattacks is to disable or switch branches from working state to off state. This, in turn, can cause an overload and a series of successive (cascading) system shutdowns. Thus, after investigating the most important peaks and fixing the types of attacks that may be attacked by the system, it is necessary to measure the changes in the system's power supply parameters (load, percentage of lost MW, etc.). Then, from the data obtained, the damage level can be calculated, and the strongest combinations of attack determined. Fig. 3 illustrates a situation where simultaneous switching off of lines {3-4, 3-5, 3-10} will cause cascade shutdown of lines {3-5, 10-7, 10-6, 4-8, 4-9} (shown by a dashed line).

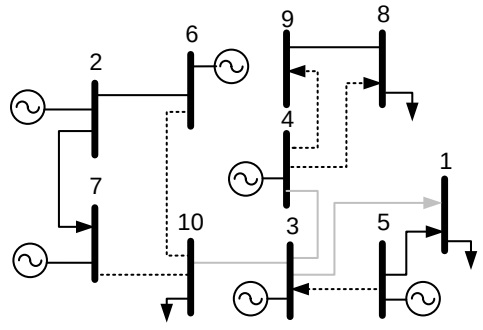

Figure 3 – Scheme of cascade shutdowns for branches 3-4, 3-5, 3-10.

A fragment of data with variants of cascading disconnections for the topological scheme shown in Fig. 1 is presented in Table VI.

TABLE VI. DATA SNIPPET WITH CASCADE SHUTDOWN OPTIONS

No	Initial combination	Cascade combination
1	2-6, 2-7	6-10, 7-10, 10-3
2	2-6, 7-10	6-10, 2-7, 10-3
3	3-4, 3-10	3-1, 3-5, 10-6, 10-7, 4-8, 4-9
4	4-9, 6-10	4-8, 4-3, 10-3, 10-7, 6-2
5	3-4, 3-5, 3-10	3-5, 10-7, 10-6, 4-8, 4-9
..	...	...

As can be seen from the Table VI, one of the key points of this analysis is that from the point of view of the attacker, it is enough to overload several lines to cause maximum damage. For example, a nozzle to cause a cascade shutdown of branches 3-1, 3-5, 10-6, 10-7, 4-8, 4-9 is enough to turn off two lines 3-4, 3-10, instead of a set of 3-4, 3- 5, 3-10.

More detailed analysis, using flow directions, gives another interesting observation - a line overloaded during a particular line cutoff (N-1) may not be overloaded for a combination of previous line cutouts. But it is obvious that N-2 outages are more harmful to the system than N-1 when considering their overall impact.

VII. CONCLUSIONS

The results of testing the method determined that one of the key points of the proposed approach is that from the point of view of the attacker, it is enough to overload several lines to cause maximum damage to the system. A more detailed analysis, using the flow directions, provided another observation - a line overloaded during a particular line cutoff (N-1) may not be overloaded for a combination of previous line cutoffs. However, it has been confirmed that N-2 outages are more harmful to the system than N-1 when considering their overall impact.

The simulation results allow us to conclude that the proposed structures can be used to find and identify the strongest combinations of attacks capable of causing the maximum system damage. To solve higher-level decision-making problems, such as obtaining realistic data on changes to system characteristics in the presence of cyber-interference, it is useful to use another type of model that directly describes physical processes in the system.

ACKNOWLEDGMENT

This research work is based upon the concept of the SPEAR project that has received funding from the European Union Horizon 2020 Research and Innovation programme under the Grant Agreement No. 787011 (SPEAR).

REFERENCES

[1] A. Greenberg, “The untold story of NotPetya, the Most Devastating Cyberattack in History”, Wired, August 2018, https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/

[2] Annex II. Security aspects of the smart grid. Smart Grid Security, 2012.

[3] P. Eder-Neuhauser, T. Zseby, J. Fabini, “Malware propagation in smart grid networks: metrics, simulation and comparison of three malware types”, Journal of Computer Virology and Hacking Techniques, 2018.

[4] M.K. Chavan, P.V. Madane, “Modelling and detection of camouflaging worms – a survey”, Int. J. Emerg. Technol. Adv. Eng., 2012, vol. 2(10), pp. 564-569.

[5] D. Moore, C. Shannon, J. Brown “Code-red: a case study on the spread and victims of an internet worm”, Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement. ACM SIGCOMM/USENIX Internet Measurement Workshop, Marseille, France, 2002, pp. 273–284.

[6] P. Li, M. Salour, X. Su, “A survey of internet worm detection and containment”, Commun. Surv. Tutor., 2008, vol. 10(1), 20-35.

[7] G. Riley, M. Sharif, W. Lee, “Simulating internet worms”, The IEEE Computer Society’s 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, 2004. (MASCOTS 2004). Proceedings. pp. 268–274.

[8] A. Dainotti, A. King, K. Claffy, F. Papale, A. Pescape, “Analysis of a “/0” stealth scan from a botnet”, IEEE/ACM Trans. Netw., 2015, vol. 23(2), pp. 341-354.

[9] J. Faulhaber, D. Felstead, H. Henry, J. Jones, “Microsoft Security Intelligence Report – Zeroing in on Malware Propagation Methods”, 2011.

[10] Regin: Top-Tier Espionage Tool Enables Stealthy Surveillance, White Paper, Symantec, 2014.

[11] B. Bencsáth, G. Pék, L. Buttyán, M. Félegyházi, “The Cousins of Stuxnet: Duqu, Flame, and Gauss. Future Internet”, 2012, vol. 4(4), pp. 971-1003.

[12] Symantec Security Response: W32.Sality | Symantec (2016). Available: https://www.symantec.com/security_response/writeup.jsp?docid=2006-011714-3948-99

[13] Kaspersky Labs: Equation Group: Questions and Answers. White Paper Report No. 1.5 #EquationAPT, Moscow, 2015.

[14] Idiom: Binaryforest—AlienSpy Java Rat Overview (2016). Available: http://blog.idiom.ca/2015/03/alienspy-java-rat-overview.html

[15] Kamluk V., Gostev A. Adwind – A Cross Plattform RAT. White Paper V. 3.0 #Adwind, Kaspersky Labs, 2016.

[16] GreyEnergy: наследник BlackEnergy атакует предприятия энергосектора. ESSET Available: https://www.esetnod32.ru/upload/iblock/211/18.10.2018-GreyEnergy-_-naslednik-BlackEnergy-atakuet-predpriyatiya-energosektora.pdf (дата звернення: 3.12.2018).

[17] N. Falliere, L.O. Murchu, E. Chien, “W32. Stuxnet Dossier”, White Paper, Security Response, Symantec, 2011, vol. 5, p. 6.

[18] N.K. Popli, A. Girghar, “Behavioural analysis of recent rasomwares and prediction of future attacks by polymorphic and metamorphic rasomware”, in Computational Intelligence: Theories, applications and Future Directions – Vol. II. Verma, Nishchal K., Ghosh, A. K. (Eds.) 2019.

[19] J. Landry, N. Izraeli, U. Shamir, C. Fenton, I. Liba “Dissecting NotPetya”, Available: https://go.sentinelone.com/rs/327-MNM-087/images/Dissecting%20NotPetya%20-%20WP.pdf?aliId=21202168

[20] R. McElroy, “Threat Report: 5 Facts to be aware of for your business”, 05 Jul 2018 Available: https://www.genusys.com/threat-report-5-facts-to-be-aware-of-for-your-business/

[21] Petya-Like Malware Campaign June 29, 2017 CERT-EU Security Advisory CERT-EU-SA2017-014: Petya-Like Malware Campaign. Available: http://cert.europa.eu/static/SecurityAdvisories/2017/CERT-EU-SA2017-014.pdf

[22] I. Ilacsu “New GreyEnergy Malware Targets ICS, Tied with BlackEnergy and TeleBots”, Available: https://www.bleepingcomputer.com/news/security/new-greyenergy-malware-targets-ics-tied-with-blackenergy-and-telebots/

[23] S. Brin and L. Page, “The anatomy of a large-scale hypertextual Web search engine”, Comput. Networks ISDN 30, 1998, pp. 107–117. Available: http://infolab.stanford.edu/pub/papers/google.pdf

[24] L. Kirichenko, D. Radivilova, A. Carlsson, “Detecting cyber threats through social network analysis: short survey”, 2018, pp. 20-34. Available: https://www.researchgate.net/publication/325215587_Detecting_cyber_threats_through_social_network_analysis_short_survey

[25] S. Paul, Z. Ni, “Vulnerability Analysis for Simultaneous Attack in Smart Grid Security”, Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT), 2017, pp. 1-5.

[26] A. Boukhtouta, D. Mouheb, M. Debbabi, O. Alfandi, E. Iqbal, M. E. Barachi “Graph-theoretic characterization of cyber-threat infrastructures”. Digital Investigation 14 (2015) S3-S15. Available: http://crossmark.crossref.org/dialog/?doi=10.1016/j.diin.2015.05.002&domain=pdf

[27] E. Hossain et al., “Application of Big Data and Machine Learning in SG, and Associated Security Concerns”. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8625421

[28] Y.Y. Yao, “Granular computing: basic issues and possible solutions”. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.64.8302&rep=rep1&type=pdf

[29] Fuzzy sets as a user-centric processing framework of granular computing. Handbook of Granular Computing. / W. Pedrycz, A. Skowron, V. Kreinovich (Eds). John Wiley & Sons Ltd, 2008. 1150 p.

[30] D. Pynes “Graphs & paths: PageRank”, Available: https://medium.com/nuances-of-programming/%D0%B3%D1%80%D0%B0%D1%84%D1%8B-%D0%B8-%D0%BF%D1%83%D1%82%D0%B8-pagerank-51193d67710d (дата звернення: 2.02.2018).

[31] S. Poudel, Z. Ni, W. Sun, “Electrical Distance Approach for Searching Vulnerable Branches During Contingencies”, IEEE Transactions on Smart Grid, 2018, vol. 9(4), pp. 3373–3382. doi:10.1109/tsg.2016.2631622.