A Brief Analysis of the 2017-2021 IEEE Energy Internet and Energy Systems Integration Conference Topics

Boris Chigarev. ORCID: 0000-0001-9903-2800

Annotation

This article is devoted to a bibliometric analysis of IEEE Energy Internet and Energy Systems Integration Conference proceedings posted in IEEE Xplore in 2017-2021. The main objective of the study was to identify the actual research issues on the Energy Internet topic. Author Keywords, INSPEC Controlled Terms, and terms compiled from n-grams derived from conference papers title and abstract texts were used to describe the actual issues. The terms were clustered using VOSviewer, and the terms included in each cluster were treated as a description of the actual issues. Brief descriptions of several publications relevant to the issue in question were given for task detail. The actual issue, which in my opinion deserves further, more detailed analysis, is proposed.

Keywords: Energy Internet, Energy Systems, IEEE Xplore, bibliometric analysis, actual research issues, VOSviewer.

Introduction

Major international specialty conferences are of particular interest for identifying actual issues within a particular topic. The topics of such conferences are carefully selected by their organizers, without which it is difficult to attract sponsors, industry representatives, and strong speakers to the conference. Conference proceedings are presented with homogeneous bibliometric data, which greatly simplifies their analysis. Selection of publications for relevance to the subject is carried out with the involvement of experts, which is more effective than the selection of publications, by keywords or filters when collecting materials for queries to abstract databases. Of course, special issues of journals have similar features, but it is difficult to imagine that a special issue will publish several hundred articles. In addition, at conferences the authors often try to announce their new works, and in special issues they more often place well-chosen articles revealing successes in solving a particular problem in general.

Institute of Electrical and Electronics Engineers (IEEE), which, according to Wikipedia is the world's largest association of engineers, with more than 423,000 members in over 160 countries, regularly conducts international conferences on issues relevant to their industry and places their bibliometric data on IEEE Xplore platform for public access.

Due to the global trend toward energy transition, the complexity of energy systems increases, they become more and more multicomponent, diverse in technical and operational characteristics. Hence, the relevance of the task of maintaining the stable operation of such complex energy systems to changes in weather conditions, accidents of its individual components, and changes in consumption and generation grows.

Historically, the analogue of such a complex distributed system whose resilience has been considered since its inception has been the Internet. This stimulated interest in the use of Internet resilience approaches developed in the course of the Internet's operation to ensure the resilience of complex distributed energy systems. That is why interest in the topic of the Energy Internet arose.

Given the growing interest in the topic of the Energy Internet, the IEEE has organized the annual IEEE Conference on Energy Internet and Energy System Integration (EI2) since 2017.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2022 doi:10.20944/preprints202211.0425.v1

According to IEEE Xplore, this conference contains the maximum number of publications among the 2021 periodicals containing the term Energy System in their title:

- 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)→797 publications
- 2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA)→262
- CSEE Journal of Power and Energy Systems→224
- Journal of Modern Power Systems and Clean Energy→177
- 2021 IEEE International Conference on Modern Electrical and Energy Systems (MEES)→167

This provoked interest in conducting a bibliometric analysis of the proceedings of this conference in order to identify actual research issues on “Energy Internet and Energy Systems Integration”.

This study uses bibliometric data from the 2017-2021 conferences. At the time the study started on November 1, 2022, data for the 2022 conference were not yet available; it was scheduled for November 11-13, 2022.

Materials and methods

The initial bibliometric data used for analysis on identifying actual research issues were the conference proceedings metadata exported from the IEEE Xplore platform in CSV format, which distribution by year was as follows: 2021→797; 2020→813; 2019→563; 2018→759; 2017→566 bibliometric records. Total 3,498 records for ("Publication Title": Conference on Energy Internet and Energy System Integration)

The aim of this study was not to perform a comprehensive bibliometric analysis of the topic in question; issues of co-authorship, affiliation of authors, distribution of publications by country, and other tasks were beyond the scope of this article.

The co-occurrence network of terms and their clustering was created using VOSviewer [1].

The data slices for more detailed analysis of the particular components of the network were built using the data exported from VOSviewer.

The advantage of the bibliometric data of this conference was that the fill rate of the Author Keywords field was over 97%, which allowed using the terms of this field to build a co-occurrence network of terms and their clustering to identify actual issues within the topic under consideration.

The following fields were used similarly: IEEE Terms, INSPEC Controlled Terms and INSPEC Non-Controlled Terms.

Additionally, a textual analysis of the "Document title" and "Abstract" fields was performed to detect key terms.

Since VOSviewer does not directly import data from IEEE Xplore platform, the fields needed for work were converted to the format of the fields as if they were exported from the Scopus system.

The total number of records used in the work was 3,478 (the previously specified value of 3498 included explanatory fields).

Document Title; Authors; Publication Year; Abstract; DOI; Author Keywords; IEEE Terms; INSPEC Controlled Terms; INSPEC Non-Controlled Terms; Article Citation Count are the names of the fields stored for work.

Similar Scopus fields: Authors; Title; Year; Cited by; DOI; Abstract; Author Keywords; Index Keywords.

doi:10.20944/preprints202211.0425.v1

IEEE Terms; INSPEC Controlled Terms; INSPEC Non-Controlled Terms and terms getting by Title and Abstract text mining were used as Index Keywords from Scopus.

Results and discussion

Author Keywords

A total of 9,722 author keywords were identified, of which 1,641 occur two or more times. The 500 terms with the highest total link strength were chosen to build the co-occurrence network; with a larger sample, the co-occurrence map of keywords is difficult to see in the graph posted in the publication; with a smaller sample, the network and its clustering may not be sufficiently detailed.

If we keep the minimum cluster size as one, we get 21 clusters. This indicates that the authors assign keywords subjectively, i.e., the very number of occurrence of keywords is large, but their co-occurrence is low.

With a minimum number of terms in a cluster being 60 there are 5 clusters, the same number of clusters is obtained with a minimum of 55 and 65 terms in a cluster. Therefore, this construction is stable. The resulting network of five clusters is shown in Figure 1.

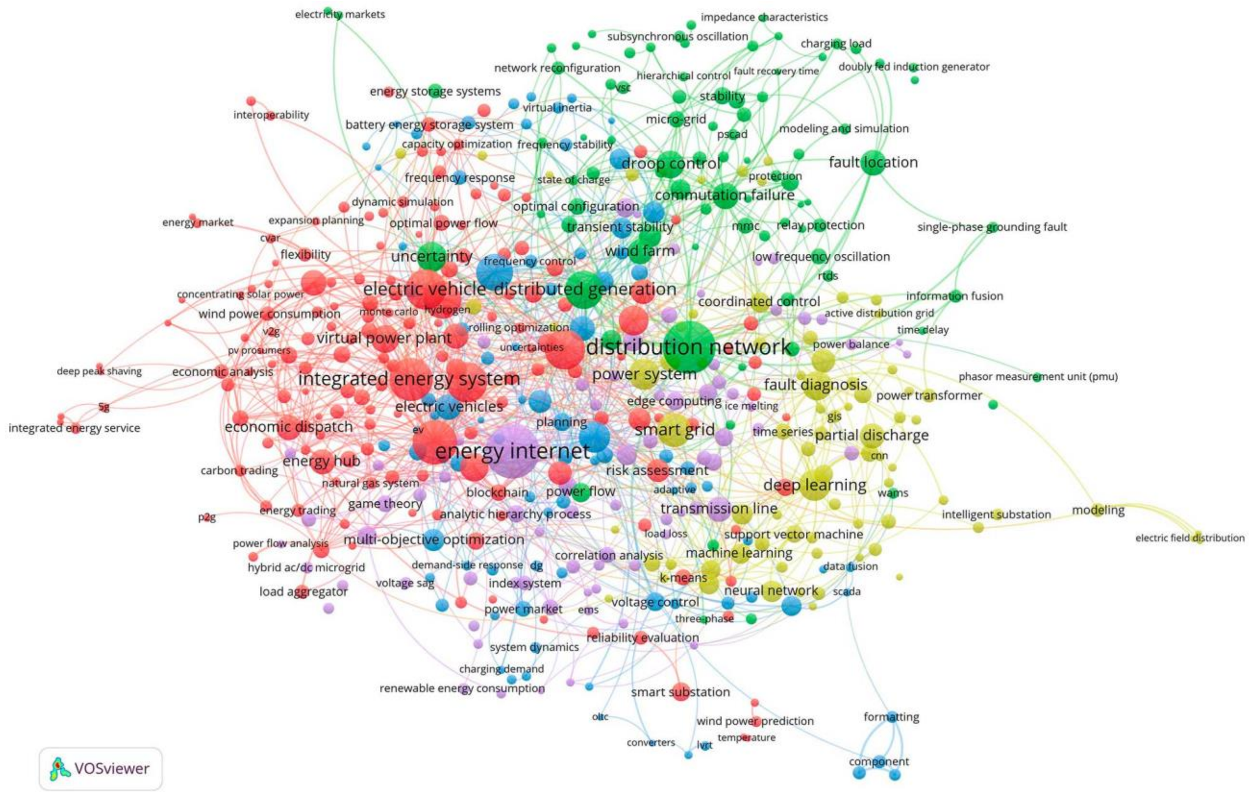

Fig.1. Five clusters of Author Keywords co-occurrence graph for 3,498 records

The variation in the occurrence of Author Keywords over time is shown in Figure 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2022 doi:10.20944/preprints202211.0425.v1

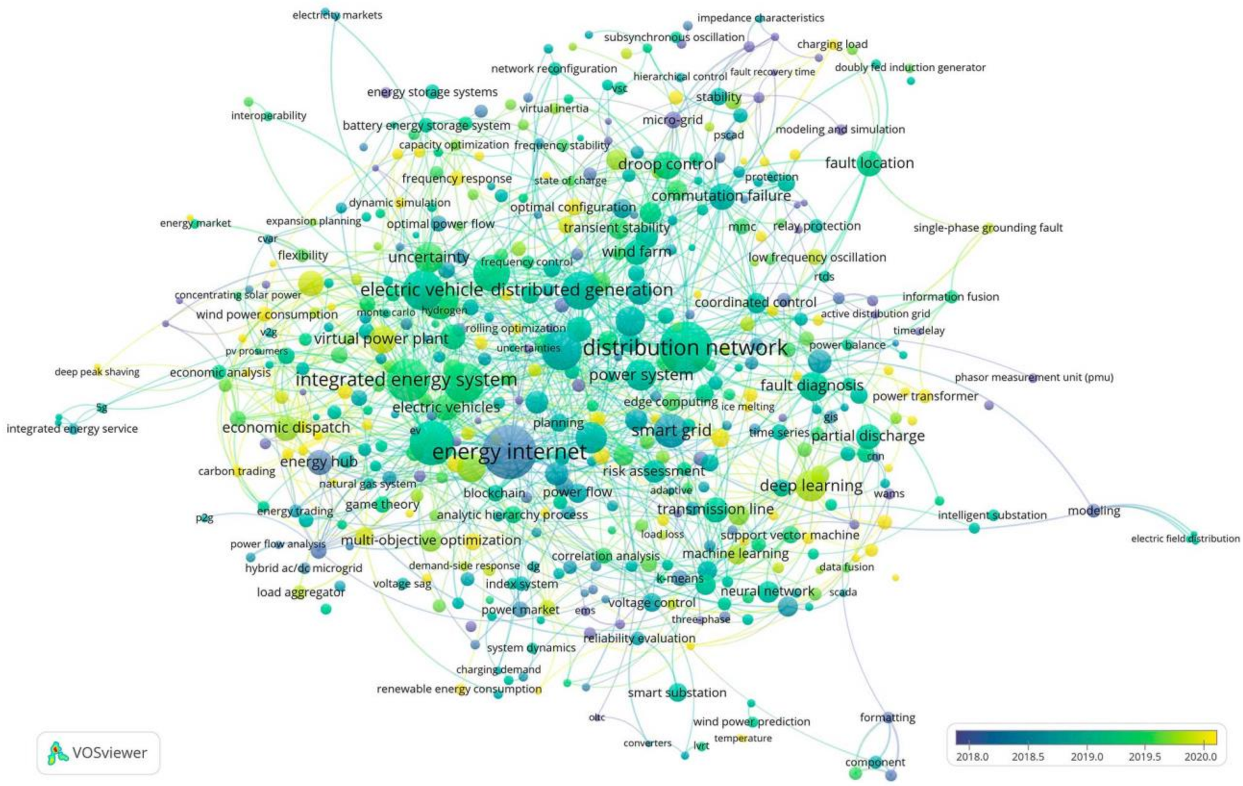

Figure 2. Temporal variation of the Author Keywords co-occurrence network for 3,498 records

The temporal variation of the Author Keywords shows that the term Energy Internet itself dominates the entire period from 2017 to 2021 and is not new.

Next, let's briefly analyze Author Keywords in each cluster.

10 most frequent Author Keywords in the **first cluster**:

demand response: 74; integrated energy system: 71; microgrid: 62; renewable energy: 62; electric vehicle: 61; wind power: 54; HVDC: 34; electricity market: 32; virtual power plant: 30; robust optimization: 26.

This cluster topic can be described as: demand response electricity market based on integrated energy system, microgrid, virtual power plant and robust optimization.

Renewable energy (renewable energy, wind power) and its applications (electric vehicle) are in the focus of publications whose keywords are included in this cluster. It is worth paying attention to HVDC - the increasing role of power grids make the tasks of their economic efficiency actual.

To assess the time shift in research interest, let's compare the 10 terms with the highest average time of occurrence and the 10 with the earlier average time of occurrence. The results are shown in Table 1.

Table 1. 10 terms of the first cluster with maximal Avg. pub. Year and minimum Avg. pub. Year.

Max avg. pub. year terms	Min avg. pub. year terms
clearing model	vsc-mtdc
hydrogen storage system	profit distribution
shared energy storage	scenario reduction
dynamic simulation	two-stage optimization
hydrogen	power-to-gas
deep peak shaving	distributed energy system
frequency regulation market	natural gas network
high renewable penetration	uncertainties

pumped-storage hydropower stations	wind power integration
reinforcement learning	super capacitor

The combination of clearing model and reinforcement learning caught my attention in the first list.

The electricity market is a typical market with imperfect competition, where electricity suppliers earn higher profits through strategic bidding behavior. To help electricity suppliers bid with limited information, a modified continuous automata algorithm with reinforcement learning is proposed in [2]. Briefly, during this process, electricity suppliers and consumers submit their bids, and the independent system operator (ISO) solves the economic dispatch problem to calculate the volume of electricity production completing a round of the market clearing process.

10 most frequent Author Keywords in the **second cluster**:

distribution network: 104; distributed generation: 56; uncertainty: 34; droop control: 30; commutation failure: 27; fault location: 26; wind farm: 25; vsc-hvdc: 19; dfig: 16; new energy: 16.

The topic of this cluster I would define as: distribution network, distributed generation and their problems: uncertainty, droop control, commutation failure, fault location. The results are shown in Table 2.

Table 2. 10 terms of the second cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Max avg. pub. year terms	Min avg. pub. year terms
optimal configuration	three-phase
doubly-fed induction generator	time delay
charging load	weak grid
fault identification	micro-grid
power flow distribution	wams
transient overvoltage	cps
model predictive control	virtual impedance
load margin	simulation test device
low frequency oscillation	simulation test method
single-phase grounding fault	wireless communication

The terms in the first column of the table agree with the list of the most commonly used terms: optimal configuration, doubly-fed induction generator, charging load, fault identification, power flow distribution, transient overvoltage, model predictive control, and other terms that have become more common lately emphasize the importance of sustainable operation of complex systems: the actual topic of the Energy Internet is rooted in this task.

The list of the 10 most common terms in the **third cluster**:

energy storage: 51; active distribution network: 38; electric vehicles: 26; control strategy: 24; unit commitment: 21; multi-objective optimization: 20; dc microgrid: 18; photovoltaic: 17; transient stability: 17; state estimation: 16.

Note that this cluster often uses the term electric vehicles, while the first cluster uses the term electric vehicle. This is one of the problems of author keywords: different spelling of terms that are close in meaning. Of course, we can lemmatize author keywords and even use synonyms, but this will be a somewhat different story. Clusters 1 and 3 have some similar themes, but while the first one focuses on demand and market, the terms in this cluster are more engineering. The topic of this cluster I would define as: energy storage, active distribution network, electric vehicles, control strategy, unit commitment, multi-objective optimization. The results are shown in Table 3

 doi:10.20944/preprints202211.0425.v1

Table 3. 10 terms of the third cluster with maximal Avg. pub. Year and minimum Avg. pub. Year.

Max avg. pub. year terms	Min avg. pub. year terms
voltage	user satisfaction
lithium-ion battery	formatting
frequency response	load shedding
equivalent model	styling
orderly charging	dynamic consistency
parameter identification	primary frequency response
sensitivity analysis	ev
incremental distribution network	decomposition
battery energy storage	dispatch
demand-side response	oltc

Even in the types of energy storage there is a difference with the first cluster, here lithium-ion batteries and orderly charging, and in the first cluster: hydrogen storage and pumped-storage hydropower stations. That is, one problem is how to optimally charge the batteries, and the second is in what to store the energy produced by renewable sources.

A list of the 10 most common terms in the **fourth cluster**:

smart grid: 45; deep learning: 37; power system: 37; fault diagnosis: 33; partial discharge: 25; simulation: 22; neural network: 18; load forecasting: 17; artificial intelligence: 16; coordinated control: 16.

In this case, the object of research are smart grid and power system, the actual tasks are: fault diagnosis and partial discharge, and possible solution methods: load forecasting and coordinated control by artificial intelligence. Table 4 lists the terms that appear in newer and earlier publications.

Table 4. 10 terms of the fourth cluster with maximal Avg. pub. Year and minimum Avg. pub. Year.

Max avg. pub. year terms	Min avg. pub. year terms
attack detection	kalman filter
digital twin	fuzzy theory
n-1 contingency	cnn
thermal aging	consumption
thermal inertia	multi-energy
loss	clean energy
modelling	modeling
non-intrusive load monitoring	active distribution grid
carbon emissions	transient stability assessment
low-voltage distribution network	urban energy internet

The terms of the earlier period are quite conservative: Kalman filter, fuzzy theory, CNN. The terms of the newer period, in the first column, may be more interesting, for example: attack detection and non-intrusive load monitoring; as the energy system becomes more complex, the role of these factors may increase and the Energy Internet approaches will be more in demand.

A list of the 10 most common terms in the **fifth cluster**:

energy internet: 110; transmission line: 24; internet of things: 16; power quality: 16; voltage stability: 15; integrated demand response: 14; deep reinforcement learning: 13; edge computing: 13; energy router: 13; power flow calculation: 13.

In this cluster, the term Energy Internet dominates by a wide margin. The terms following it bring additional definitions to it, for example: power quality and voltage stability. But for me, the sequence of terms is of the greatest interest: integrated demand response, deep reinforcement

 doi:10.20944/preprints202211.0425.v1

learning, edge computing, which together define a very interesting topic for research on complex energy systems. Table 5 lists the terms in the fifth cluster that appear in newer and earlier publications

Table 5. 10 terms of the fifth cluster with maximal Avg. pub. Year and minimum Avg. pub. Year.

Max avg. pub. year terms	Min avg. pub. year terms
electricity retail market	electric heating
optimize operation	multi-microgrid
multiple time scales	real-time simulation
deep reinforcement learning	distributed optimization
renewable energy consumption	ems
photovoltaic power	voltage quality
extreme learning machine	reactive power
analytic hierarchy process (ahp)	energy management system
combined heat and power	cyber-attack
ice melting	cyber-physical system (cps)

In the list of newer terms, along with the objects of research: electricity retail market and renewable energy consumption, a multi-level approach to solving current problems can be seen: multiple time scales, analytic hierarchy process (ahp), deep reinforcement learning, which only clarifies the above-mentioned topics: **integrated demand response, deep reinforcement learning, edge computing**.

To explore this topic, here is a brief description of a number of publications demonstrating the implementation of this topic in particular cases.

For example, the authors of [3] argue that existing smart building management algorithms based on optimization suffer from the high cost of both building-specific modeling and on-demand computational resources. Using a surrogate building model derived automatically from building operation data, the reinforcement learning (RL) agent proposed by the authors learns the optimal management policy on the cloud infrastructure, and then this policy is extended to the edge devices for execution.

With the proliferation of smart appliances (e.g., appliances with computing and data analysis capabilities) and high-performance computing devices (e.g., graphics processing units) in the households, it is expected of surging residential energy consumption caused by computation. The integrated home energy management system (HEMS) aims to maximize the homeowner's expected total reward, defined as the reward from completing edge computing tasks minus the cost of electricity consumption, the cost of computation offloading to the cloud, and the penalty of violating the demand side management (DSM) requirements. Therefore, it is important to schedule edge computing as well as traditional energy consumption in a smart way, especially when the demand for computation and thus for electricity occurs during the peak hours of electricity consumption. This issue is analyzed in the article [4].

The existing cloud computing paradigm is stubborn to address issues and challenges such as rapid response and local autonomy. To address this challenge, the authors of [5] consider a power control framework combining edge computing and reinforcement learning, which makes full use of edge nodes to sense network state and control power equipment to achieve the goal of fast response and local autonomy. Additionally, the authors focus on the non-convergence problem of power flow calculation, and combine deep reinforcement learning and multi-agent methods to realize intelligent decisions, with designing the model such as state, action, and reward.

 doi:10.20944/preprints202211.0425.v1

The demand for improving productivity in manufacturing systems makes the industrial Internet of things (IIoT) an important research area spawned by the Internet of things (IoT). Communications between massive heterogeneous industrial devices and clouds will cause high latency and require high network bandwidth. In this article, the authors of [6] use deep reinforcement learning (DRL) to solve the scheduling problem in edge computing to improve the quality of services provided to users in IIoT applications. Next, the authors propose a hierarchical scheduling model considering the central-edge computing heterogeneous architecture. Double Deep Q-Network (DDQN) framework is proposed to make scheduling decisions for communication.

Power Internet of Things (PIoT) is a promising solution to meet the increasing electricity demand of modern cities, but real-time processing and analysis of huge data collected by the devices is challengeable due to limited computing capability of devices and long distance from the cloud center. In the paper [7], the authors consider the edge computing assisted PIoT where the computing tasks of the devices can be either processed locally by the devices, or offloaded to edge servers. Aiming to maximize the long-term system utility which is defined as a weighted sum of reduction in latency and energy consumption, the authors propose a novel task offloading algorithm based on deep reinforcement learning, which jointly optimizes task scheduling, transmit power of the PIoT devices, and computing resource allocation of the edge servers.

Due to increasing complexity, uncertainty and data dimensions in power systems, conventional methods often meet bottlenecks when attempting to solve decision and control problems. Deep reinforcement learning (DRL) is one of these data-driven methods and is regarded as real artificial intelligence (AI). This field of research has been applied to solve a wide range of complex sequential decision-making problems, including those in power systems. The overview of this topic is given in the article [8].

Bibliometric analysis and clustering of terms based on their co-occurrence depend significantly on text preprocessing and slice formation, while helping to avoid exaggerated subjectivity in the selection of promising research topics. The individual expert often works as a filter, focusing on tasks close to him.

Authors of publications can also be seen as experts, so their keywords, titles, and article abstracts reflect research priorities. Using bibliometric data from their publications not only reduces the possible bias of individual experts, but also allows us to see interesting problems that a particular expert may not have focused on. Conducting a broad survey of experts' opinions on a particular topic also reduces bias, but is time-consuming and costly and may not be available to a small group of researchers.

INSPEC Controlled Terms (Keywords assigned to articles from a controlled vocabulary of over 10,000 scientific terms created by INSPEC[1])

All 3,498 records contain 1,723 INSPEC Controlled Terms; 556 of them meet threshold 5, 1,066 terms meet threshold 2.

As in the previous case, we use the 500 terms with the highest number of links to build the co-occurrence network. The reason for choosing a threshold of 2 terms is that it is better to use a larger sample to select the most linked terms.

The number of INSPEC Controlled Terms in our records is significantly lower than the Author Keywords, 1,723 versus 9,722. More than half of them occur two or more times (1,060), while for Author Keywords, 1,641 terms out of 9,722 meet this criterion. The large co-

---

[1] https://ieeexplore.ieee.org/Xplorehelp/searching-ieee-xplore/command-search

occurrence of INSPEC Controlled Terms and their small diversity led to a small number of clusters (4), even if the parameter of the minimum number of terms in the cluster is equal to one.

The picture of the co-occurrence network is compact, see Fig. 3, and the change in term occurrence over time is minimal, see Fig. 4.

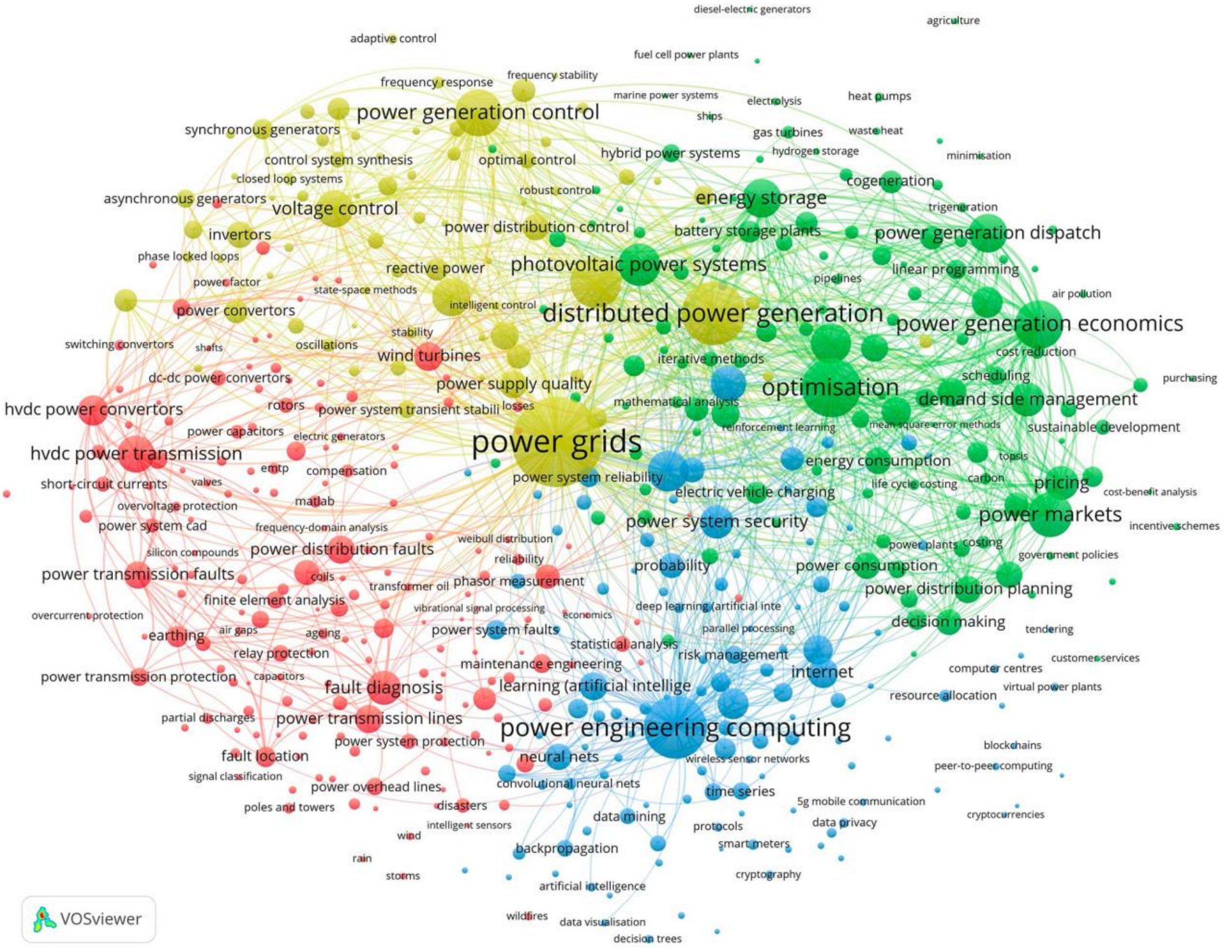

Figure 3. INSPEC Controlled Terms co-occurrence map for 3,498 records

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2022 doi:10.20944/preprints202211.0425.v1

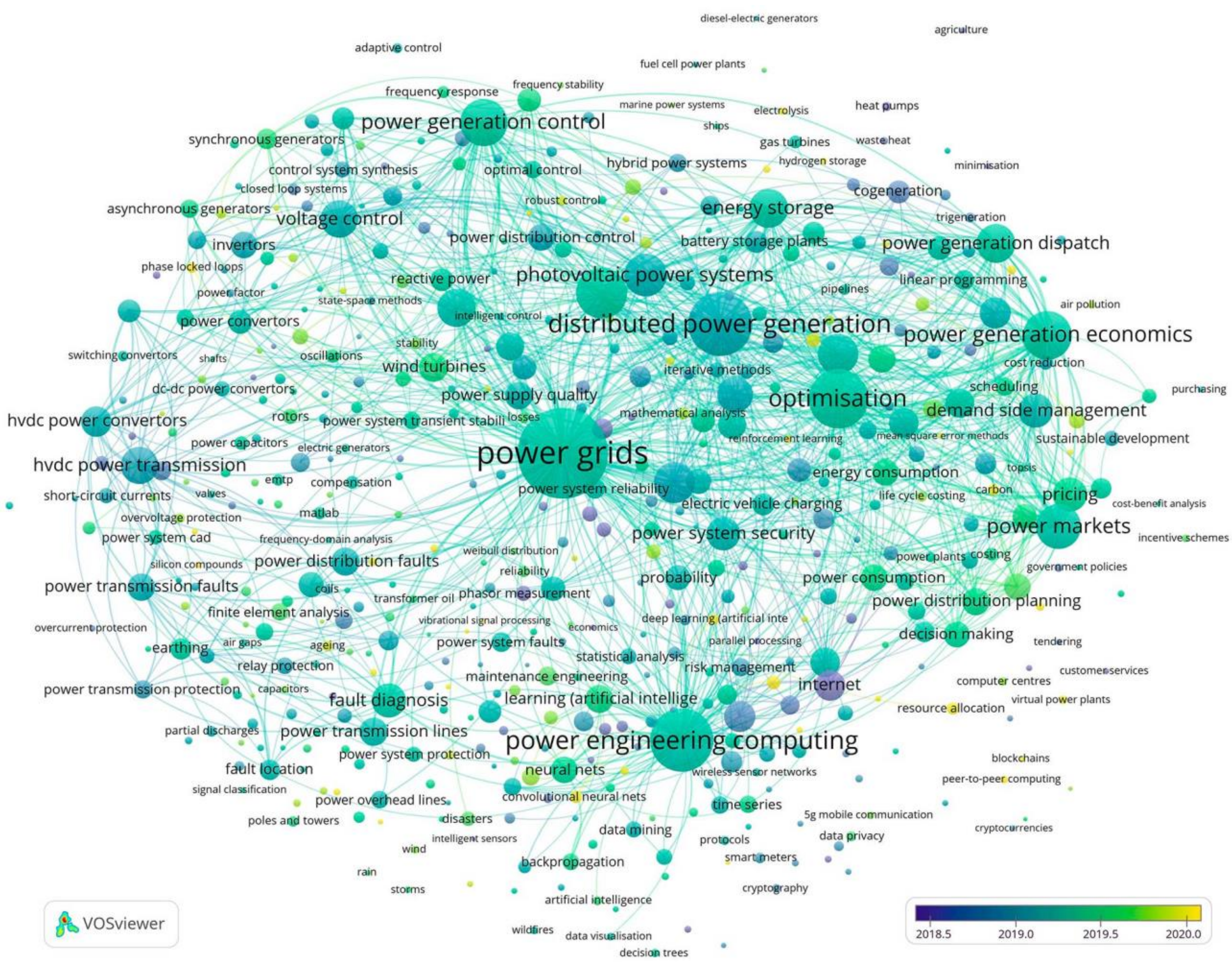

Figure 4. Change over time of the INSPEC Controlled Terms co-occurrence network for 3,498 records

Visually, we can note that the colors of the main terms refer to the middle of the time range, which is significantly different from Fig. 2.

It should in addition noted that the key term in the query to the IEEE Xplore abstract database "Energy Internet" is not observed among the INSPEC Controlled Terms.

Projecting the topic of the publication onto an established vocabulary leads, on the one hand, to well-defined clusters; on the other hand, it is impossible to identify newly utilized terms.

Thus, using INSPEC Controlled Terms to identify new topics of interest for more detailed consideration, as was done for Author Keywords, is not advisable. The INSPEC Controlled Terms are of interest as filters when refining queries to a reference database, for example, if we want to restrict the sample to tasks related to "power engineering computing".

INSPEC Non-Controlled Terms

All 3,498 records include 3,168 INSPEC Non-Controlled Terms; 618 of these meet threshold 5, 3,441 terms meet threshold 2.

Using the 500 terms with the highest number of relationships that meet threshold 2 to construct a network of terms co-occurrence, we get 5 clusters by using minimum 65 terms in cluster.

By definition INSPEC Non-Controlled Terms is "Additional keywords assigned to articles which describe the topics or subjects of a document. These terms are not part of the INSPEC controlled vocabulary and include new and emerging concepts". Therefore, it is reasonable to compare these terms with INSPEC Controlled Terms and identify those that do not occur in the controlled vocabulary. The comparison is made for the 500 terms from each list with the highest total number of links, i.e., those used to build the term's co-occurrence network.

Only 58 of the 500 terms compared are in the overlap. This means that these are significantly different samples.

Below is a list of 10 terms that appear frequently in 500 INSPEC Non-Controlled Terms but do not appear in 500 INSPEC Controlled Terms, checked by direct search to eliminate variations in the spelling of terms such as energy system or energy systems: energy internet — 96; integrated energy system — 62; distributed generation — 56; energy storage system — 45; control strategy — 38; smart grid — 38; active distribution network — 33; demand response — 32; electricity market — 26; microgrid — 23.

These terms are not like newly appearing terms. To check this, I searched for these terms in the headings and annotations for each year. The results are shown in Table 6.

Table 6: Term frequency by year in the texts of titles and abstracts

Term	17N	17R	18N	18R	19N	19R	20N	20R	21N	21R
energy internet	178	78	138	55	48	23	70	31	49	20
integrated energy system	44	22	73	34	33	12	76	29	129	52
distributed generation	58	35	53	31	19	12	51	25	64	30
energy storage system	54	31	48	29	44	25	76	35	72	41
control strategy	121	67	130	66	72	41	85	50	141	69
smart grid	47	26	42	29	16	10	43	21	23	15
active distribution network	27	13	36	17	5	2	38	15	30	17
demand response	49	24	55	23	34	15	56	28	115	43
electricity market	13	12	23	14	24	12	34	18	49	23
microgrid	177	44	188	50	84	27	128	36	170	50

Note: The format of the headings in the table is: 17-21 (2017-2021) years, N — the number of times the term occurs in the titles and annotations of a given year, R — the number of records in which the term occurs.

Table 6 shows that these terms appear in conference publication titles and abstracts throughout the years. Therefore, INSPEC Non-Controlled Terms should not be used as pointers to emerging topics.

The terms "INSPEC Controlled Terms" and "INSPEC Non-Controlled Terms" are far from the words found in the texts of titles and abstracts, they rather reflect the categories to which the system assigns a particular entry in IEEE Xplore. Therefore, to identify promising targets, it is advisable to identify key terms and analyze their co-occurrence in the texts of titles and abstracts of conference proceedings, as was done for Author Keywords.

The simplest way to find key terms in texts is to identify n-grams. This approach requires text preprocessing of titles and annotations. In this case, the text was lowercase, then the stop words were removed, followed by stemming according to the Krovetz method.

N-grams as Key Terms (N-gram Terms)

Titles and annotations of 3498 records gave me 16754 N-gram Terms of which 2261 occur 5 or more times.

By default, the 500 N-gram Terms are grouped into 6 clusters (only 6 terms in cluster 6), including at least 10 terms in a cluster will result in 5 clusters. These five clusters are shown in Figure 5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2022 doi:10.20944/preprints202211.0425.v1

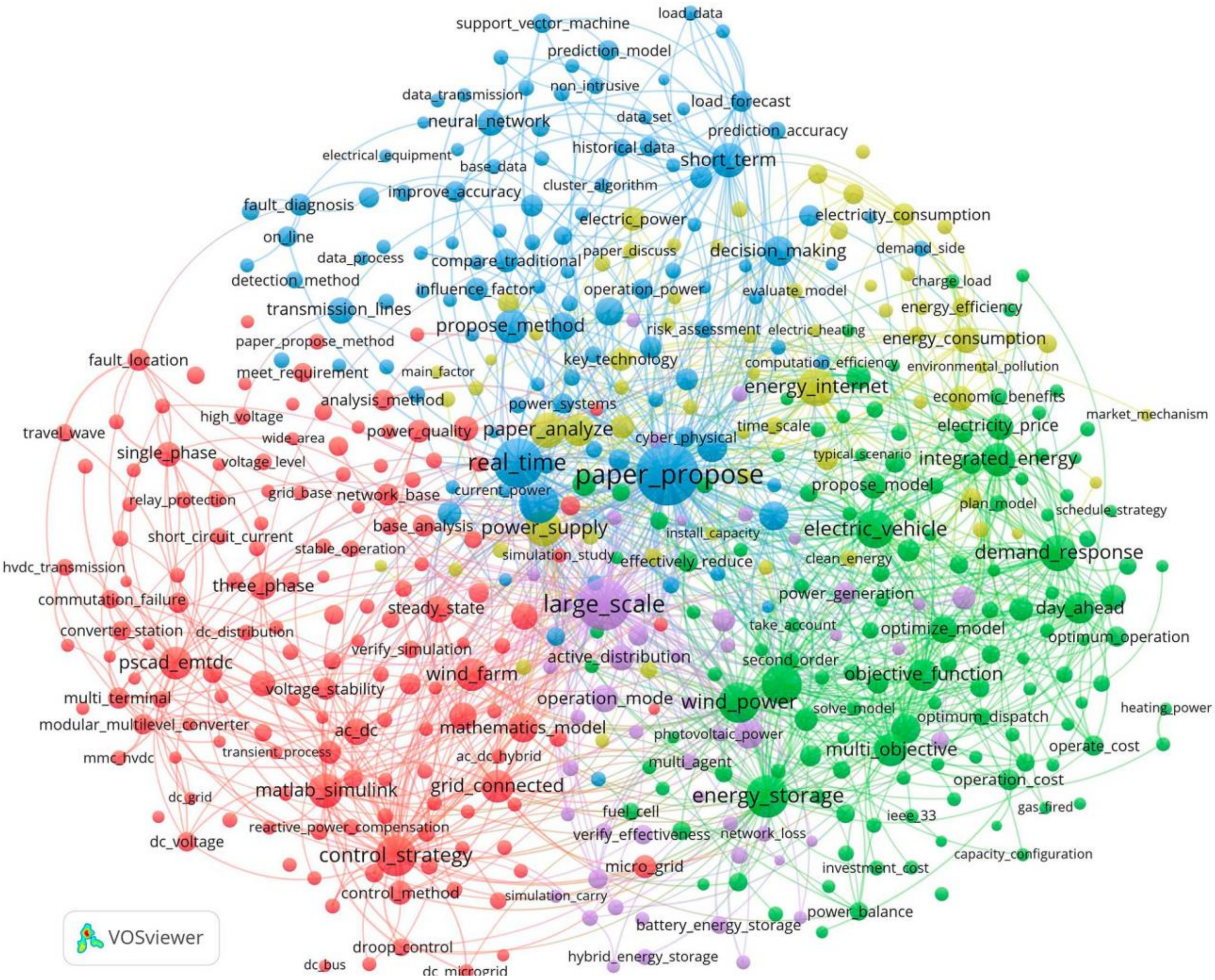

Figure 5. Co-occurrence map of N-gram Terms from Titles and Abstracts fields for 3,498 records

By analogy with Author Keywords, for each cluster we construct tables containing 10 N-gram Terms with the maximum average publication time and the minimum average publication time. For the **first cluster**, the results are given in Table 7.

Table 7. 10 terms of the first cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Max avg. pub. year terms	Min avg. pub. year terms
low_voltage_distribution	operation_control
transient_overvoltage	micro_grid
calculation_model	control_mode
circuit_breaker	bus_voltage
reliability_power_supply	theoretical_basis
voltage_direct_current	droop_control
optimum_control	output_power
wind_turbines	fault_conditions
verify_simulation	wide_area
doubly_fed_induction_generator	hvdc_transmission

The data obtained for the N-gram Terms are significantly different from the Author Keywords, they are more specialized, while the Author Keywords are more general in nature.

In order to reveal the possible topics of this cluster, it is useful to give examples of publications in which the terms listed in the first column of the table meet.

 doi:10.20944/preprints202211.0425.v1

In [9] the authors proposed a calculation model applicable to multilayer feedforward networks in which weighted summation with positive and negative weights is performed separately in each layer, and the summation results are then fed to the next layers without subtraction operation. Simulation results showed that the energy efficiency for calculating weighted sums was 290~TOPS/W, which is more than an order of magnitude higher than in modern digital AI processors, even though the minimum interconnect width used in the PoC chip was several times larger than in such digital processors.

The paper [10] proposes a time-domain method to calculate the fault response in multiterminal dc (MTdc) grids and the performance of hybrid dc breakers. Based on the analytical results, three parameters of the hybrid dc circuit breaker, i.e., current limiting reactor, arrester rated voltage, and time delay are optimally selected with respect to maximum overcurrent, maximum overvoltage, fault clearance time, and energy absorption in arresters through multiobjective optimization. The proposed method, based on traveling waves: 1) provides a sound representation of fault performance by considering all created traveling waves; 2) introduces a new approach to estimate the reflection coefficients; and 3) provides an approximation of the worst-case fault location.

Microgrids (MG) are well described as distribution systems with integrated distributed energy resources, i.e., photovoltaic (PV) and wind power generation systems. In the paper [11], an efficient system for controlling the maximum power of a microgrid fed by PV modules and a wind turbine is proposed. For the proposed PV maximum power point tracking (MPPT), the boost chopper duty cycle is being adjusted to set the PV panel operating point to a maximum power point. Whereas, for the proposed wind turbine, a slope angle controller is designed to force the wind turbine to achieve optimum operation in low to medium wind speeds.

As can be seen from the given descriptions of the three articles, they refer to very specific tasks. Therefore, the n-gram terms we have chosen should mainly be used by subject matter experts as hints for making queries to abstract databases in accordance with their professional interests.

Let's do the same routine for the terms of the **second cluster** and present the results in Table 8.

Table 8. 10 terms of the second cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Max avg. pub. year terms	Min avg. pub. year terms
total_cost	energy_supply
economic_operation	power_demand
energy_conversion	combine_cooling
capacity_configuration	natural_gas
particle_swarm_optimize_algorithm	energy_demand
low_carbon	charge_load
wind_power_consumption	optimum_solution
optimum_configuration	security_constrained
life_cycle_cost	energy_flow
key_issue	energy_hub

As you can see from the table, newer and older terms can be associated with an economic task. Older terms are less specific.

The following are examples of articles in which terms from the first column of the table occur.

The articles were taken as examples from various publications; it is the topic, not the source, that is important.

The article [12] deals with the production of electricity using renewable energy microgrid (REM) that is a prerequisite for achieving one of the Sustainable Development Goals (SDG 7 - Affordable and Clean Energy). The microgrid configuration proposed in the paper minimized CO2 emissions (by 92.3%) and fuel consumption (by 92.4%) compared to using a fossil-fueled diesel generator. Determination of the optimal REM size is associated with several nonconvexities and nonlinearities, which excludes the use of deterministic optimization search methods.

A hybrid wind-solar PV hybrid microgrid (MG) with biomass and energy storage is proposed in [13] to meet the load demand for buildings in the intended region and provide economic dispatch as well as power trading by supplying/receiving energy to/from the utility grid. A Hybrid Grey Wolf with Cuckoo Search Optimization (GWCSO) was used to achieve the optimal size of the proposed MG connected to the grid. The control operations plan included storage batteries to compensate for energy shortages if priority resources (wind turbine and solar PV) were unable to meet demand. Renewable energy systems, especially in countries with limited fossil fuel resources, are promising and environmentally sustainable sources of electricity generation.

Similarly, to the previous one, the table 9 for the **third cluster** terms is made up.

Table 9. 10 terms of the second cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Max avg. pub. year terms	Min avg. pub. year terms
model_power	power_load
data_driven	method_propose
safe_operation_power	monte_carlo_method
convolution_neural_network	k_means
weather_conditions	data_transmission
load_data	smart_grid
current_power	communication_technology
prediction_accuracy	base_data
influence_factor	power_fluctuate
electrical_equipment	computation_efficiency

The common impression is that this n-grams cluster refers to data processing methods for energy systems: model_power, convolution_neural_network, prediction_accuracy, influence_factor, monte_carlo_method, k_means and computation_efficiency.

The following are examples of articles in which terms from the table occur.

Forecasting demand in the power system is the most important task in the field of power engineering. It is important to build reliable and efficient forecasting models to ensure accurate load forecasting. In [14], three methods are used for short-term load forecasting: deep neural network (DNN), multilayer perceptron-based artificial neural network (ANN), and decision tree (DR) forecasting. The results show that the DNN model outperforms the other models and is statistically different from them.

The review article [15] identifies and articulates a research problem related to the field of load prediction. Accurate modeling and sophisticated short-term load forecasting (STLF) analysis are becoming increasingly important for microgrids (MGs). For a quick overview of the results, they are presented in the form of tables.

Now repeat the same procedure for the terms in the fourth cluster and show the results in Table 10.

Table 10. 10 terms of the fourth cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2022 doi:10.20944/preprints202211.0425.v1

Max avg. pub. year terms	Min avg. pub. year terms
carbon_emissions	energy_saving
spot_market	business_model
analyze_impact	electric_power
install_capacity	energy_management
comprehensive_evaluate	propose_base
renewable_energy_consumption	evaluate_method
power_source	environmental_pollution
main_factor	development_energy_internet
high_quality	interconnecte_power
quantitative_evaluate	energy_internet

It follows from this table that neither energy_internet nor energy_saving are new terms in this cluster. And Carbon_emissions, spot_market, analyze_impact - these terms are receiving more and more attention in publications.

Here are some typical articles for the topic: carbon_emissions, spot_market, analyze_impact and electric_power, energy_management.

The carbon trading market is an important policy tool that contributes to China's carbon neutrality goals. The study [16] focuses on identifying the relationship between the carbon market and other related markets. The results show that the information spillover effect between China's pilot carbon markets, the energy market and the stock market is relatively low. For investors and policymakers, looking at each market from a systems perspective will provide a more accurate picture of them and their interconnections.

To achieve carbon neutrality by 2050, Japan should accelerate the reduction of its dependence on fossil fuels, especially in energy-intensive sectors. One way is to implement a carbon pricing system that converts emissions from fossil fuels into production and consumption costs. A study [17] examines the correlation between the price in the wholesale electricity spot market and the cost of carbon in nine regions of Japan through the carbon cost transfer ratio.

Repeating the same procedure for the terms of the fifth cluster, we get the results shown in Table 11.

Table 11. 10 terms of the fifth cluster with maximal Avg. pub. Year and minimum Avg. pub. Year

Max avg. pub. year terms	Min avg. pub. year terms
frequency_modulation	network_topology
pump_storage	solar_energy
peak_shaving	distribute_energy
voltage_deviation	distribute_generator
frequency_regulation	strategy_base
virtual_power_plant	multi_agent
grid_connection	simulation_carry
distribute_photovoltaic	power_output
battery_energy_storage	renewable_energy_resources
energy_storage_system	distribute_power_supply

In this cluster the total number of terms is less (50) than in the previous ones (147 in the first one), so they look more coherent in the two columns. However, among the new terms are more those related to energy storage: pump_storage, battery_energy_storage, energy_storage_system.. Accordingly, examples of publications found on these terms are as next:

The electric power industry is undergoing a paradigm shift toward a carbon-free smart system, aided by growing energy demand, deterioration of long-lived physical assets, and global

 doi:10.20944/preprints202211.0425.v1

environmental concerns. Energy storage systems (ESS) can help maximize these opportunities and mitigate potential problems. The paper [18] focuses on the ESSs installed in end-user premises and the corresponding technologies, the various billing and pricing policies, and their potential opportunities from the perspective of system operators and end-users.

The Cloud Energy Storage System (CES) is a shared distributed energy storage resource. The random disordered charging and discharging of large-scale distributed energy storage has a large impact on the energy system. The paper [19] focuses on two objectives: to present detailed plans for designing an ordered managed CES system in a realistic power system and to analyze the load curves of five types of distributed energy storage systems using Monte Carlo simulation (MCS) to manage and operate the CES system.

Figure 6 shows the overall change in the occurrence of terms over time.

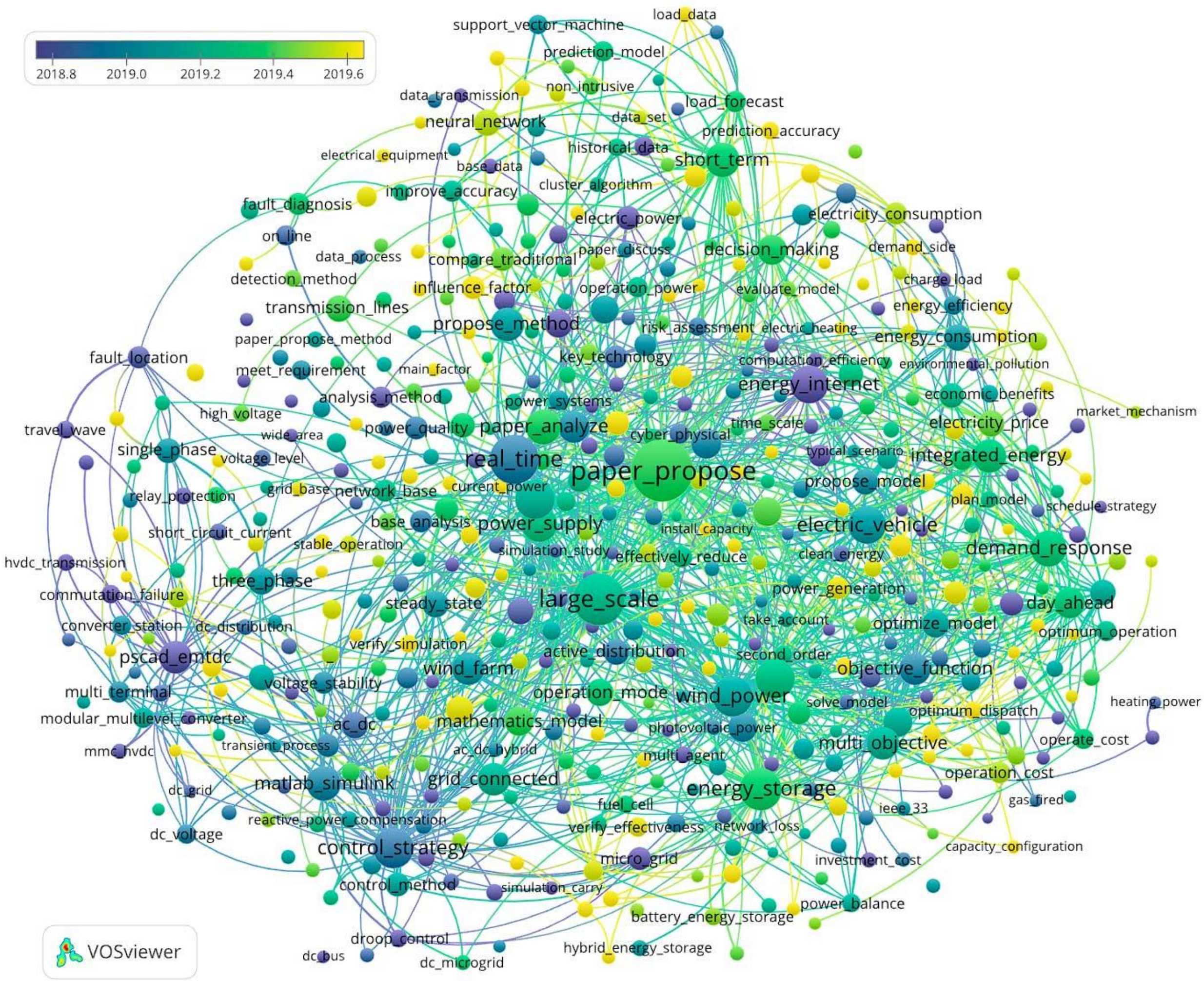

Figure 6. Changes over time in the co-occurrence network of n-grams from Titles and Abstracts fields for 3,498 records.

Comparing this figure with Figure 4, we see that the terms expressed as n-grams from Titles and Abstracts change significantly more over time than INSPEC Controlled Terms. Consequently, they are more appropriate for identifying emerging topics.

Conclusions

1. Bibliometric analysis of scientific publication texts to identify relevant research issues is often used as part of a systematic review, for which it is necessary to reasonably select several dozen articles from the thousands published during the period under review. Based on the above analysis, I propose the following steps for the selection of publications, the full texts of which will be used to find answers to the questions posed in the systematic review (recommendations are given for IEEE Xplore):

    1. Use INSPEC Controlled Terms to create broad filters for selecting publications suitable for review.
    2. The resulting sample is then narrowed with Author Keywords as an expert opinion on the attribution of the article to the analyzed issue. Author words are more specific than INSPEC Controlled Terms, they are less numerous, and they change more often over time (which is important for assessing trends).
    3. In the third step, it is advisable to identify frequently occurring terms in the titles and abstracts of publications (e.g., on compiling a list of n-grams) that will allow the subject specialist compiling a systematic literature review to select terms that will help find answers to the questions posed in the review in the full texts of publications.

2. As a result of this analysis, I hold the opinion that the topic “A comprehensive demand-side power market based on an integrated energy system including microgrids, using the virtual power plant concept and optimization techniques using deep reinforcement learning and edge computing”, is an actual issue for further, more detailed research.

References

[1] N. J. van Eck and L. Waltman, “Software survey: VOSviewer, a computer program for bibliometric mapping,” *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010, doi: 10.1007/s11192-009-0146-3.

[2] Q. Jia, Y. Li, Z. Yan, C. Xu, and S. Chen, “A Reinforcement-learning-based Bidding Strategy for Power Suppliers with Limited Information,” *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 4, pp. 1032–1039, 2022, doi: 10.35833/MPCE.2020.000495.

[3] X. Zhang, D. Biagioni, M. Cai, P. Graf, and S. Rahman, “An Edge-Cloud Integrated Solution for Buildings Demand Response Using Reinforcement Learning,” *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 420–431, Jan. 2021, doi: 10.1109/TSG.2020.3014055.

[4] T. Li, Y. Xiao, and L. Song, “Deep Reinforcement Learning Based Residential Demand Side Management With Edge Computing,” in *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Beijing, China, Oct. 2019, pp. 1–6. doi: 10.1109/SmartGridComm.2019.8909778.

[5] T. Pu *et al.*, “Power flow adjustment for smart microgrid based on edge computing and multi-agent deep reinforcement learning,” *J Cloud Comp*, vol. 10, no. 1, p. 48, Dec. 2021, doi: 10.1186/s13677-021-00259-1.

[6] J. Wu, G. Zhang, J. Nie, Y. Peng, and Y. Zhang, “Deep Reinforcement Learning for Scheduling in an Edge Computing-Based Industrial Internet of Things,” *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–12, Aug. 2021, doi: 10.1155/2021/8017334.

[7] J. Hu, Y. Li, G. Zhao, B. Xu, Y. Ni, and H. Zhao, “Deep Reinforcement Learning for Task Offloading in Edge Computing Assisted Power IoT,” *IEEE Access*, vol. 9, pp. 93892–93901, 2021, doi: 10.1109/ACCESS.2021.3092381.

[8] “Deep reinforcement learning for power system: An overview,” *CSEE JPES*, 2019, doi:

10.17775/CSEEJPES.2019.00920.
[9] Q. Wang, H. Tamukoh, and T. Morie, “A Time-domain Analog Weighted-sum Calculation Model for Extremely Low Power VLSI Implementation of Multi-layer Neural Networks,” 2018, doi: 10.48550/ARXIV.1810.06819.
[10] Y. Song, J. Sun, M. Saeedifard, S. Ji, L. Zhu, and A. P. S. Meliopoulos, “Optimum Selection of Circuit Breaker Parameters Based on Analytical Calculation of Overcurrent and Overvoltage in Multiterminal HVDC Grids,” *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 4133–4143, May 2020, doi: 10.1109/TIE.2019.2921279.
[11] M. M. A. Mahfouz, M. Alsumiri, and R. Althomali, “Efficient Power Utilization Control Scheme for Hybrid Distribution Generation Grid,” *JEEE*, vol. 9, no. 1, p. 26, 2021, doi: 10.11648/j.jeee.20210901.14.
[12] A. L. Bukar, C. W. Tan, D. M. Said, A. M. Dobi, R. Ayop, and A. Alsharif, “Energy management strategy and capacity planning of an autonomous microgrid: Performance comparison of metaheuristic optimization searching techniques,” *Renewable Energy Focus*, vol. 40, pp. 48–66, Mar. 2022, doi: 10.1016/j.ref.2021.11.004.
[13] A. M. Jasim, B. H. Jasim, and V. Bureš, “A novel grid-connected microgrid energy management system with optimal sizing using hybrid grey wolf and cuckoo search optimization algorithm,” *Front. Energy Res.*, vol. 10, p. 960141, Sep. 2022, doi: 10.3389/fenrg.2022.960141.
[14] M. A. Alotaibi, “Machine Learning Approach for Short-Term Load Forecasting Using Deep Neural Network,” *Energies*, vol. 15, no. 17, p. 6261, Aug. 2022, doi: 10.3390/en15176261.
[15] V. Y. Kondaiah, B. Saravanan, P. Sanjeevikumar, and B. Khan, “A review on short-term load forecasting models for micro-grid application,” *The Journal of Engineering*, vol. 2022, no. 7, pp. 665–689, Jul. 2022, doi: 10.1049/tje2.12151.
[16] Y. Yao, L. Tian, and G. Cao, “The Information Spillover among the Carbon Market, Energy Market, and Stock Market: A Case Study of China’s Pilot Carbon Markets,” *Sustainability*, vol. 14, no. 8, p. 4479, Apr. 2022, doi: 10.3390/su14084479.
[17] D. Ding, “The impacts of carbon pricing on the electricity market in Japan,” *Humanit Soc Sci Commun*, vol. 9, no. 1, p. 353, Oct. 2022, doi: 10.1057/s41599-022-01360-9.
[18] M. Rezaeimozafar, R. F. D. Monaghan, E. Barrett, and M. Duffy, “A review of behind-the-meter energy storage systems in smart grids,” *Renewable and Sustainable Energy Reviews*, vol. 164, p. 112573, Aug. 2022, doi: 10.1016/j.rser.2022.112573.
[19] J. Li, Y. Xing, and D. Zhang, “Planning Method and Principles of the Cloud Energy Storage Applied in the Power Grid Based on Charging and Discharging Load Model for Distributed Energy Storage Devices,” *Processes*, vol. 10, no. 2, p. 194, Jan. 2022, doi: 10.3390/pr10020194.