

ChatGPT 4 ▾



You

selected_papers.xlsx
Spreadsheet

I am sending you an Excel file containing the titles and abstracts of 36 academic articles focused on cyber risk. Review the provided articles and categorize them into distinct groups based on their focus. Assign a descriptive name to each group, reflecting the common theme of its articles. List the IDs and titles of the articles in each group. Additionally, explain the rationale behind the groupings, focusing on how the content or perspective on cyber risk unifies the articles within each group. Ensure every article is categorized. Describe the logic or algorithm used for this classification.



ChatGPT

I have successfully loaded the Excel file containing the titles and abstracts of 36 academic articles focused on cyber risk. The data includes an ID, title, and abstract for each article.

To categorize these articles, I will use a content-based classification approach. This involves analyzing the titles and abstracts to identify key themes, topics, or methods that are common across multiple articles. The main steps in this process are:

1. **Text Preprocessing:** Clean and preprocess the text data from titles and abstracts to facilitate analysis. This includes removing special characters, converting text to lowercase, and possibly tokenizing the text.
2. **Keyword Extraction:** Identify significant keywords or phrases from the titles and abstracts. This helps in understanding the focus of each article.
3. **Topic Modeling or Clustering:** Use natural language processing techniques such as topic modeling (e.g., LDA) or clustering algorithms (e.g., K-means) to group articles based on the similarity of their content.
4. **Group Naming and Categorization:** Analyze the common themes within each group to assign a descriptive name that reflects the collective focus of the articles in that group.
5. **Rationale Explanation:** Provide explanations for each group based on the common themes or perspectives on cyber risk that unify the articles within the group.

I will start by preprocessing the text and then proceed with the classification. Let's begin with the preprocessing step. [↗]
The articles have been initially categorized into five topics based on their content. This categorization was achieved using Latent Dirichlet Allocation (LDA), a popular topic modeling technique. LDA identifies topics by discovering abstract "topics" that occur in a collection of documents.

Now, I will examine the top keywords for each topic to identify the common theme and assign a descriptive name to each group. Additionally, I will list the IDs and titles of the articles in each group, along with an explanation of the rationale behind these groupings. Let's proceed with identifying the top keywords for each topic. [↗]

Based on the top keywords for each topic, the articles can be categorized into distinct groups as follows:

1. **Cloud and Mobile Security:**
 - **Keywords:** cloud, cyber, security, care, mobile, products, quality, health, insurance, imbalanced
 - **Rationale:** This group focuses on cyber security issues in cloud computing and mobile environments. It includes discussions on security in healthcare technology and the quality of cyber products.
2. **Malware and Critical Infrastructure Protection:**
 - **Keywords:** risk, ci (critical infrastructure), cyber, malware, security, families, medical, institutions, telemedicine, improve
 - **Rationale:** Articles in this group are centered around risks associated with malware and protecting critical infrastructures, particularly in the medical and telemedicine sectors.
3. **Advanced Detection Techniques in Cybersecurity:**
 - **Keywords:** cyber, learning, detection, insurance, cybersecurity, deep, feature, risk, transportation, network
 - **Rationale:** This category includes articles that discuss advanced detection techniques in cybersecurity, such as deep learning and feature detection, applied in various domains including transportation and network security.
4. **Cyber Risks in IoT and Organizational Contexts:**
 - **Keywords:** cyber, insurance, risk, security, IoT (Internet of Things), devices, risks, organizations, audits, market



- **Rationale:** The focus here is on cyber risks associated with IoT devices and their impact on organizations. It includes discussions on audits and market considerations related to these risks.
5. **Data Security and Communication in Cyber Defense:**
- **Keywords:** data, climate, security, communication, information, corporate, capability, defense, intrusion, neural
 - **Rationale:** Articles in this group explore data security and communication strategies in the context of cyber defense, including corporate information security and intrusion detection systems.

Next, I will list the IDs and titles of the articles in each group. [~]

The articles have been categorized into five distinct groups based on their focus and themes related to cyber risk. Here are the groups along with the articles they contain:

1. **Group 1: Cloud and Mobile Security**

- **Articles:**
 - ID 3: Cyberattack detection in mobile cloud computing: A deep learning approach
 - ID 6: Deep and Machine Learning Approaches for Anomaly-Based Intrusion Detection of Imbalanced Network Traffic
 - ID 16: Towards a maturity model for health-care cloud security
 - ID 19: Using Risk Assessments to Assess Insurability in the Context of Cyber Insurance
 - ID 22: Cyber-Warranties as a Quality Signal for Information Security Products
 - ID 34: Preventing the Drop in Security Investments for Non-competitive Cyber-Insurance Market

2. **Group 2: Malware and Critical Infrastructure Protection**

- **Articles:**
 - ID 1: DeepOrigin: End-to-End Deep Learning for Detection of New Malware Families
 - ID 13: Risk management-based security evaluation model for telemedicine systems
 - ID 14: Evaluation of Mobile Health Applications: Is Regulatory Policy Up to the Challenge?
 - ID 33: A framework for incorporating insurance in critical infrastructure cyber risk strategies
 - ID 35: Optimal Security Configuration for Cyber Insurance
 - ID 36: The evolution of cyber-insurance industry and market: An institutional analysis

3. **Group 3: Advanced Detection Techniques in Cybersecurity**

- **Articles:**
 - ID 2: Autoencoder-based feature learning for cyber security applications
 - ID 4: Toward an Online Anomaly Intrusion Detection System Based on Deep Learning
 - ID 5: Network intrusion detection for cyber security using unsupervised deep learning approaches
 - ID 8: A Self-Adaptive Deep Learning-Based System for Anomaly Detection in 5G Networks
 - ID 11: Phishing in healthcare organisations: threats, mitigation and approaches
 - ID 12: Clinical Cybersecurity Training Through Novel High-Fidelity Simulations
 - ID 18: RiskWriter: Predicting Cyber Risk of an Enterprise
 - ID 21: Incentive Contract for Cybersecurity Information Sharing Considering Monitoring Signals
 - ID 25: An insurance theory based optimal cyber-insurance contract against moral hazard
 - ID 28: Cyber risk and insurance for transportation infrastructure
 - ID 31: SECONDO: A Platform for Cybersecurity Investments and Cyber Insurance Decisions

4. **Group 4: Cyber Risks in IoT and Organizational Contexts**

- **Articles:**
 - ID 10: Infotech. Cyber Security. Health Care Learns to Share Scares and Solutions
 - ID 15: Cyber risk cost and management in IoT devices-linked health insurance
 - ID 17: Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem
 - ID 20: Differentiating Cyber Risk of Insurance Customers: The Insurance Company Perspective
 - ID 24: What are the actual costs of cyber risk events?
 - ID 26: Cyber-Insurance as a Signaling Game: Self-reporting and External Security Audits
 - ID 27: Post-incident audits on cyber insurance discounts
 - ID 29: Cyber Insurance Market in Israel – What is the Official Policy?
 - ID 32: Three Layer Game Theoretic Decision Framework for Cyber-Investment and Cyber-Insurance

5. **Group 5: Data Security and Communication in Cyber Defense**

- **Articles:**
 - ID 7: A Signal Processing Approach for Cyber Data Classification with Deep Neural Networks
 - ID 9: Information security climate and the assessment of information security risk among healthcare employees
 - ID 23: A framework for effective corporate communication after cyber security incidents
 - ID 30: A Cybersecurity Insurance Model for Power System Reliability Considering Optimal Defense Resource Allocation

These groupings were made based on the common themes or perspectives on cyber risk that unify the articles within each group, as indicated by their titles, abstracts, and the identified keywords using topic modeling techniques [32].