



Unveiling 5G Core Research Trends: Thematic Clustering and LLM-Enhanced Insights into Architecture, Security, and Artificial Intelligence

Hüseyin Parmaksız^{1*}, Harun Şeker²

¹ Department of Management Information Systems, Faculty of Economics and Administrative Sciences, Bilecik Şeyh Edebali University, Bilecik, Türkiye

² Department of Computer Engineering, Graduate Education Institute, Bilecik Şeyh Edebali University, Bilecik, Türkiye

huseyin.parmaksiz@bilecik.edu.tr, harun@harunseker.org

Abstract

5G Core is the backbone of digital transformation with ultra-low latency, network slicing, and dense IoT integration. In the study, 635 academic studies indexed in Web of Science were clustered using Leiden, Louvain, HDBSCAN, and K-Means, and five main themes were identified: service-based architecture, security (post-quantum, blockchain), artificial intelligence/machine learning Network Data Analytics Functionality (NWDAF), deep reinforcement learning), cloud/edge computing, and network resource management. The clustering methods were evaluated using the Silhouette, Davies-Bouldin, and Calinski-Harabasz quality metrics. Although K-means appears to be at the top in terms of quality metrics, Leiden and Louvain outperform K-means in terms of interpretability. A hybrid Large Language Models (LLMs) approach was used for theme interpretation. General themes were identified using Qwen3-Max-Preview, and then the previously identified themes were supported using cloud-based free Nemotron, gpt-oss-20b, and Sonoma SkyAlpha models via OpenRouter. Qwen3 stands out in terms of semantic depth and summarization accuracy. This methodology, which combines quantitative clustering with qualitative LLM analysis, highlights the strategic role of 5G Core in both communication and as the architecture, security, and AI infrastructure of digital transformation, in a multidimensional way.

Keywords: 5G core network, Large language models, Artificial intelligence, Cyber security

5G Çekirdek Araştırma Trendlerinin Açıklanması: Tematik Kümeleme ve Mimari, Güvenlik ve Yapay Zeka Konularında LLM ile Güçlendirilmiş İlgörüler

Öz

5G Core, ultra düşük gecikme süresi, ağ dilimleme ve yoğun IoT entegrasyonu ile dijital dönüşümün temel bileşenlerindedir. Bu çalışmada Web of Science veritabanında "5G Core" ifadesi aranmış ve bulunan 635 akademik yayın Leiden, Louvain, HDBSCAN ve K-Means kümeleme algoritmaları kullanılarak analiz edilmiştir. Analiz sonucunda; hizmet tabanlı mimari, güvenlik (kuantum sonrası, blok zinciri), yapay zeka/makine öğrenimi (NWDAF, derin pekiştirme öğrenimi), bulut/kenar bilişim ve ağ kaynak yönetimi olmak üzere beş ana tema belirlenmiştir. Kümeleme performansı Silhouette katsayısı, Davies-Bouldin indeksi ve Calinski-Harabasz skoru gibi içsel kalite metrikleri kullanılarak değerlendirilmiştir. Nicel metrikler açısından K-Means algoritması en yüksek performansı göstermiş olsa da, Leiden ve Louvain yöntemlerinin ortaya çıkardığı kümelerin tematik tutarlılığı ve yorumlanabilirliği daha yüksek bulunmuştur. Bu nedenle, nicel kalite ölçütlerinin ötesinde, kümelerin anlamsal anlamlandırılabilirliği de değerlendirmeye dâhil edilmiştir. Tematik yorumlama aşamasında hibrit bir Büyük Dil Modelleri (LLMs) yaklaşımı benimsenmiştir. İlk olarak Qwen3-Max-Preview modeli kullanılarak genel tematik yapı çıkarılmış, ardından OpenRouter platformu üzerinden erişilen Nemotron, gpt-oss-20b ve Sonoma SkyAlpha modelleri ile bu temalar doğrulanmış ve zenginleştirilmiştir. Karşılaştırmalı analizler, Qwen3 modelinin anlamsal derinlik, kavramsal tutarlılık ve özetleme doğruluğu bakımından daha üstün performans sergilediğini göstermiştir. Nicel kümeleme ile nitel LLM analizini birleştiren bu metodoloji, 5G Core'un sadece iletişimde değil, aynı zamanda çok boyutlu bir şekilde dijital dönüşümün mimarisi, güvenliği ve yapay zeka altyapısı olarak da stratejik rolünü ortaya koymaktadır.

Anahtar kelimeler: 5G çekirdek ağı, Büyük dil modelleri, Yapay zeka, Siber güvenlik

* Corresponding Author.

E-mail: huseyin.parmaksiz@bilecik.edu.tr

Received : 22 Sep 2025

Revision : 15 Dec 2025

Accepted : 19 Jan 2025

1. Introduction

Fifth-generation (5G) mobile networks are not just a radio access technology, but a network paradigm representing a radical transformation in the entire communication infrastructure. At the heart of this transformation is a 5G Core Network (5GC) built on software-defined (SDN), virtualized network function (NFV), and cloud-native principles, replacing the traditional hardware-based, monolithic, and rigid Evolved Packet Core (EPC) architecture. 5GC is defined by the Service-Based Architecture (SBA), introduced with 3rd Generation Partnership Project (3GPP) Release 15 and continuously developed with Releases 16 and 17. This architecture transforms network functions (NFs: AMF, SMF, UPF, NRF, UDM, AUSF, NWDAF, NEF, etc.) into independent, reusable microservices, providing operators with unprecedented flexibility, scalability, and speed compared to previous generations. The strategic importance of 5G is not limited to technical specifications (eMBB, URLLC, mMTC) (Popowski et al., 2018). 5G, by enabling the creation of multiple, isolated, end-to-end virtual networks on a single physical infrastructure through network slicing, makes it possible to simultaneously meet the unique and contrasting service needs of vertical sectors such as automotive, manufacturing, smart cities, telemedicine, and energy (e.g., 1 ms latency or millions of device connections). This capability transforms 5GC from a mere communication channel into the cornerstone of the digital economy and society. However, this revolutionary flexibility and software-defined nature have also brought new and complex challenges in areas such as network management, security, and performance optimization. The separation of control and user planes (CUPS), HTTP/2-based service interfaces (SBI), and end-to-end slicing management require rethinking traditional security models and network management approaches. In this context, 5GC is poised to become a self-managing (autonomous) and proactive system powered by artificial intelligence (AI) and machine learning (ML).

This study rigorously analyzes the overlooked reliance on intrinsic clustering metrics as proxies for thematic validity. While indices such as the Silhouette Coefficient, Calinski–Harabasz, and Davies–Bouldin exhibit mathematical rigor, they are insensitive to semantic nuances, favoring geometric compactness and separation over conceptual coherence. In rapidly progressing technical domains like 5G Core, where terminology evolves more swiftly than established vocabularies (NWDAF, AIaaS, zero-touch orchestration, etc.), a cluster with a high CH score may mistakenly conflate fundamentally distinct research paradigms (post-quantum cryptography and blockchain-based network slicing, etc.) due to superficial lexical similarities. To alleviate this

epistemic risk, we propose a hybrid validation protocol that integrates three complementary dimensions: (i) geometric quality (as a necessary yet insufficient baseline), (ii) intra-cluster semantic consistency (measured through NPMI on top-k keywords), and (iii) LLM-enhanced interpretability (utilizing various free-tier models, including Nemotron, gpt-oss-20b, and Sonoma SkyAlpha via OpenRouter, to generate, compare, and cross-validate thematic narratives). This multi-axis paradigm promotes reproducibility and aligns clustering results with the goal and discursive logic of academic work, transforming clusters from mere statistical artifacts into epistemically grounded knowledge units. Leveraging this methodological framework, the study seeks to elucidate the research trends, structural development, and prospective advancements of 5GC across its architectural, security, and intelligence facets by systematically thematically clustering 635 academic studies utilizing Leiden, Louvain, HDBSCAN, and K-means, while enhancing the findings through profound semantic analysis facilitated by a hybrid LLM-augmented validation pipeline.

2. Related Works

Through a comprehensive literature analysis, three main research dimensions of 5G core networks have been identified: Architecture and Virtualization, Security and Privacy, and Integration with Artificial Intelligence and Autonomous Management. These categories highlight key research directions and summarize the common findings of 635 studies.

Architecture and Virtualization: The Cloud Native Revolution and Resilience Engineering Microservices and Service-Oriented Architecture (SOA): SBA constitutes the cornerstone of 5GC (Wang et al., 2022). RESTful APIs provide communication amongst NFTs. This provides fault tolerance, scalability, and deployment autonomy. Recent research indicates that SBA is not solely for the control plane; it is also crucial for 6G since it integrates with the user plane (UPF) and the radio access network (RAN). By categorizing NFs based on processes, methodologies like "Procedure-Based Stateless Design" (PP5GS) aim to significantly reduce resource utilization and signaling traffic (Akashi et al., 2025).

Kubernetes and other container orchestration technologies are compatible with 5G Core in a cloud-native and containerized environment. This facilitates resource efficiency, high availability (HA), and dynamic scaling of Network Function Virtualizations (NFVs) (Ma et al., 2018). A fundamental trend for adaptable resource distribution and rapid recovery in mistake situations is "stateless" network function design.

Control and User Plane Separation (CUPS): Distributing the User Plane Function (UPF) to the network edge facilitates local data breakout and

reduces latency. Significant efforts are being directed towards the development of hardware acceleration techniques such as DPDK, eBPF, XDP, and P4 programmable switches to improve UPF performance. SmartNIC-based UPF solutions (Singh et al., 2022) provide significant performance enhancements by accelerating both data plane and control plane (PFCP) activities.

Network Slicing and Resource Management: A fundamental breakthrough of 5G is network slicing. Regulating slice admission and distributing resources is a complex optimization challenge. To resolve this issue, several AI/ML models have been suggested in the literature, including Deep Learning (LSTM, CNN), Reinforcement Learning (DRL), Federated Learning (FL), and Genetic Algorithms. These models maximize cross-slice resource allocation by predicting future traffic loads, facilitating resource sharing between flexible and non-flexible slices, and proactively adjusting resource capacity. Two prevalent techniques for automating the slice lifecycle are "Zero-Touch" automation and closed-loop management systems (Coronado et al., 2022).

Safeguarding and Confidentiality: The Zero Trust Framework and Ascendant Threats. The Security Challenges of the Small Business Administration: The HTTP/2-based SBI is vulnerable to novel attack techniques (Ming et al., 2024), including stream multiplexing assaults. Moreover, the virtualized environment introduces new vulnerabilities such as container escape and hypervisor assaults. Zero Trust architectures, which surpass traditional firewall approaches to ensure continuous verification of all network functions and devices, are receiving increased attention.

Key Agreement and Authentication (AKA): The 5G-AKA protocol enhances privacy by obscuring the subscriber's permanent identifier (SUPI). Nevertheless, due to the potential vulnerability of post-quantum cryptography (PQC) to quantum computers, research into its integration has intensified. Empirical research suggests that the integration of PQC algorithms, such as KYBER768 (Corsi et al., 2025), into 5GC has minimal impact on bandwidth and latency.

Defense against Distributed Denial of Service (DDoS) attacks: DDoS attacks are cyberattacks that intentionally overload a server, network, or server with excessive traffic, causing massive congestion and denying access to authorized users. These attacks can be carried out using a botnet comprised of compromised systems such as computers, smartphones, and Internet of Things devices. DDoS attacks primarily employ two techniques: overloading the target with packets to exhaust resources and exploiting protocol or application flaws to disrupt services with faulty packets. Some attacks indirectly target intermediate infrastructures, preventing service access (Çakmak, 2024).

Such attacks can cause serious operational disruptions, particularly in modern telecom infrastructures like the 5GC, by targeting critical components such as AMF (Access and Mobility Management Function) and UPF (User Plane Function). Therefore, as an effective defense strategy, we recommend adopting innovative approaches such as core-level packet filtering techniques (e.g., eBPF/XDP) (Sharaf et al., 2022), machine learning-assisted anomaly detection (LSTM Autoencoder, XGBoost, Transformer-based models), and "quarantine slices" designed to isolate attack traffic (Candal et al. 2020). These methods play a critical role in maintaining service continuity by ensuring both early detection of attacks and effective isolation and mitigation.

Data privacy and surveillance: 5G has the potential to track users' locations and activities. This is a significant concern for sensitive sectors such as the military and healthcare. The solutions underscore the utilization of temporary identifiers (SUCI) while minimizing the reliance on permanent identifiers such as SUPI (Tsai and Chen, 2025). To ensure transparency and auditability, blockchain technology is being examined in areas such as resource allocation and sharding management (Quan et al., 2024).

Integrating machine learning with artificial intelligence: The establishment of the Network Data Analytics Function (NWDAF) for autonomous networks: The NWDAF, as delineated by 3GPP, examines network traffic, predicts consequences, and supplies data to other network activities utilizing machine learning models. It is utilized in domains such as resource optimization, anomaly detection, and user behavior analysis. Numerous NWDAFs can collaboratively train models while preserving data privacy through the application of NWDAF in conjunction with Federated Learning (FL). According to Release 17, NWDAFs will have the capability to exchange ML models with each other (Nadar and Harri, 2024).

Autonomous Network Management: Artificial Intelligence and Machine Learning are employed to automate functions such as network slice generation, scalability, and error recovery. For example, DRL-based algorithms can optimize the distribution of network resources (Ullah et al., 2023), the strategic placement of UPFs, and the adaptive scaling of NFVs such as AMFs. This reduces operational expenditures and enhances service standards for operators.

RAN Intelligence and Open RAN (O-RAN): The O-RAN design facilitates the integration of AI into the RAN. The RAN Intelligent Controller (RIC) can enhance the RAN by employing machine learning models via xApps and rApps (Qazzaz et al., 2024). 5GC generates a comprehensive "intelligence layer." The NWDAF in 5GC can be linked with radio telemetry data from the RAN to enhance resource allocation and traffic prediction (Manias et al., 2022).

Increased demand on network resources can affect signal strength and efficiency. To overcome these challenges, anomaly detection techniques are often used to identify and predict potential network issues. At this stage, the HDBSCAN algorithm is applied to detect outliers and abnormal patterns in the data (Pawar and Deshmukh, 2025).

This study uses a dataset of 635 studies from the Web of Science to analyze the research ecosystem of the 5G Core network. It uses semantic representation, clustering, and topic structure identification to identify five main research dimensions and cross-cutting sub-themes such as Edge Computing, Network Slicing, and IoT Integration. These themes often intersect with core technical domains like Architecture, Cloud/Edge Convergence, AI/ML Applications, and Security Mechanisms. The study validates the thematic mapping's robustness and confirms the consistency of research trends. Hybrid domains like AI + Security and AI + Architecture are emerging as dominant trajectories, signaling a shift towards intelligent, adaptive, and autonomously secured network infrastructures. Thematic visualization via word clouds and LLM-based interpretive synthesis further demonstrates how advanced language models can automate scientific knowledge consolidation while mitigating human bias with key findings graphically summarized in Figure 1.

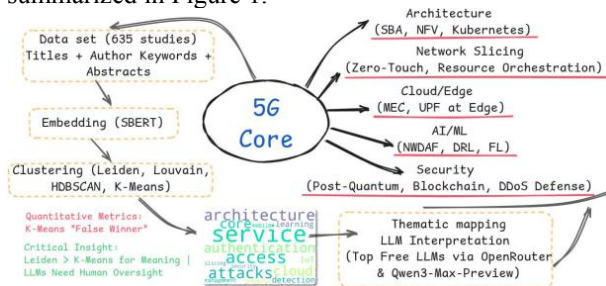


Figure 1. Graphical Abstract

3. 5G Core Network Architecture

The 5G architecture is divided into two main planes: the User Plane and the Control Plane, which organize the network's different functions. The planes separate control data, user data, and network management functions from each other. This provides a more flexible, scalable, and high-performance structure. The overall architectural layout of this functional separation is illustrated in Figure 2.

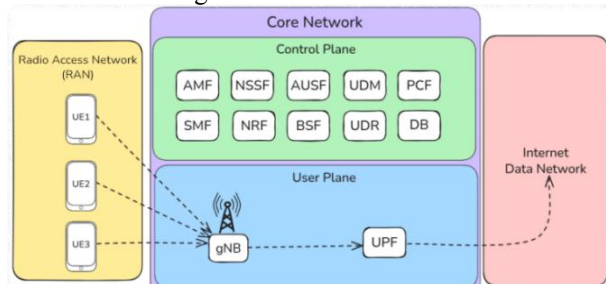


Figure 2. 5G Core Network Architecture

The 5G network architecture manages user data traffic via the User Plane, which, in conjunction with the User Plane Function (UPF) and the Next Generation NodeB (gNB), facilitates the transport, routing, and delivery of data packets to ensure uninterrupted internet and private network access for end users (Leyva-Pupo et al., 2019). The Control Plane operates autonomously while collaborating with the User Plane, augmenting network flexibility and scalability by managing essential control functions, including user authentication, session initiation and maintenance, mobility management, roaming, and the dynamic implementation of Quality of Service (QoS) policies through core components such as the Access and Mobility Management Function (AMF), Session Management Function (SMF), and Policy Control Function (PCF). User Equipment (UE), including smartphones, tablets, and IoT devices, functions as the principal access point, utilizing SIM/eSIM credentials (such as IMSI and authentication keys) to secure connectivity with the network through gNB, the radio access node that facilitates 5G New Radio (NR) for high-throughput, ultra-low latency, and mission-critical applications. In the core network, the AMF regulates access and mobility processes, whereas the SMF manages session lifespan, IP address allocation, and QoS policy implementation by routing traffic via the UPF. Authentication is performed by the Authentication Server Function (AUSF), while subscriber profiles and policy data are centrally administered by the Unified Data Management (UDM) and Unified Data Repository (UDR). The Policy Control Function (PCF) dynamically implements network policies to enhance resource allocation, whereas the Binding Support Function (BSF) preserves session-to-function binding data to guarantee consistent interactions among network functions. The Network Repository Function (NRF) supports service-based architecture by enabling the discovery and interoperability of core components, while the Network Slice Selection Function (NSSF) intelligently allocates customized network slices according to user context and service needs, thus facilitating precise resource distribution and realizing 5G's potential for personalized, efficient, and scalable service provision (Şeker and Parmaksız, 2025). These components collectively create a modular, interoperable, and security-conscious architecture that comprehensively enhances user experience, operational efficiency, and adaptive network management in next-generation mobile systems.

4. Clustering Algorithms & Quality Metrics

Clustering is a crucial unsupervised learning topic, and the clustering effect directly impacts the analysis of clustering findings. Finding an effective and accurate clustering impact evaluation technique is one

of the hardest topics that researchers have been attempting to address (Wang and Xu, 2019).

The communities with higher modularity values perform better. The values of modularity are always less than 1. When the modularity number is less than one, it indicates that each node functions as its own community. Subgraphs occur with a small number of internal edges and a large number of intercommunity connections (Anuar et. Al., 2021). The definition of modularity is given in Eq. (1) below.

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j) \quad (1)$$

where A_{ij} represents the weight of the edge between i and j , $k_i = \sum_j (A_{ij})$ is the sum of the weights of the edges attached to vertex i , C_i is the community to which vertex i is assigned, the δ function $\delta(u, v)$ is 1 if $u=v$ and 0 otherwise and $m = \frac{1}{2} \sum_{i,j} (A_{ij})$ (Blondel et al., 2008).

In 2008, Blondel et al. introduced the Louvain algorithm, a widely adopted method for community detection based on modularity optimization (Anuar et. Al., 2021). The algorithm operates in two phases. First, nodes are iteratively moved between neighboring communities to maximize modularity locally. Second, communities are aggregated into super-nodes to form a new, simplified network, enabling recursive optimization. Figure 3 depicts the iterative process of the Louvain algorithm. The algorithm starts with a singleton partition, in which each node creates its own community (a). It then moves individual nodes between communities to maximize modularity, resulting in an optimal partition (b). Based on this partition, an aggregate network is created, with each community compressed into a single node (c). On the coarsened network (d), the algorithm repeats the node-moving optimization. These processes of local optimization followed by aggregation are repeated recursively until no more improvement in modularity can be obtained.

While efficient, Louvain is limited to static networks and may produce poorly connected or arbitrarily fragmented communities (Anuar et. Al., 2021).

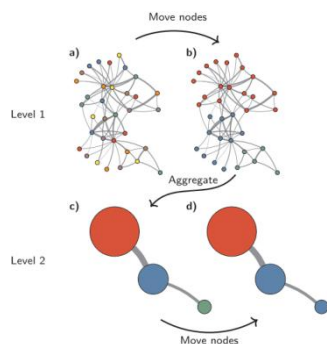


Figure 3. Infographic of Louvain algorithm (Traag et al., 2019).

In 2019, Vincent Traag et al. proposed the Leiden algorithm as a significant improvement over Louvain. Designed to overcome Louvain's structural weaknesses, Leiden guarantees that all detected communities are well-connected, while preserving and often surpassing Louvain's computational efficiency. The algorithm extends the two-phase structure with an additional intermediate step: after initial modularity optimization, it introduces a *refinement phase* that explicitly partitions communities into internally cohesive subgroups before the aggregation step. This ensures robust, topologically sound communities without sacrificing scalability. The Leiden algorithm thus comprises three distinct phases. Local modularity optimization is achieved through node movement (1). Partition refinement enforces internal connectivity (2). Network aggregation enables iterative coarsening (3). The algorithm performs effectively across small, medium, and large-scale networks. This makes it particularly suitable for complex, evolving systems.

Figure 4 depicts the step-by-step execution of the Leiden algorithm, as reported by Traag et al. (2019). The technique begins with a singleton partition (a), then relocates nodes to maximize modularity (b), and subsequently incorporates a vital refining step to ensure well-connected communities (c). An aggregate network is then built (d), in which the red community in (b) splits into two distinct subcommunities in (c), resulting in two separate nodes within the same community in (d). Node movement continues in the aggregated network (e), and if additional refining is not required, the partition remains intact (f). These repeated stages continue until convergence, ensuring structural coherence and stability of optimization.

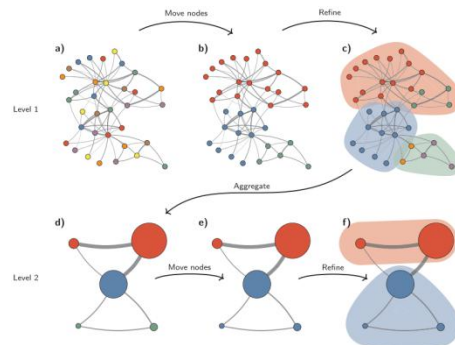


Figure 4. Infographic of Leiden algorithm (Traag et al., 2019).

In 5G Core research analysis, Leiden clustering enables the identification of thematically coherent and structurally robust research communities. This technique works within vast scientific corpora. Ensuring both internal cohesion and clear intercommunity boundaries provides a more reliable foundation for mapping knowledge domains. It also aids in tracking research trends and revealing hidden collaboration patterns in scholarly networks.

The k-means method is one of the best unsupervised clustering algorithms. It is used a lot in online user-facing applications (Manias et al., 2022).

The Silhouette index (Zhou and Gao, 2014) focuses on the cohesiveness and separation of detected clusters. Cohesion $a(x_i)$ indicates how closely connected the objects in a cluster are. It is expressed by the average distance between the element x_i and the other elements in the same cluster. Separation $b(x_i)$ quantifies how distinct or well-separated one cluster is from others. It is defined as the minimum average distance between element x_i and all other elements in any cluster. The Silhouette Coefficient (SC) is defined as follows, as shown in Eq. (2):

$$SC = \frac{1}{n} \sum_{i=1}^n s(x_i) \quad (2)$$

$$\text{where } s(x_i) = \frac{b(x_i) - a(x_i)}{\max\{a(x_i), b(x_i)\}} \quad (3)$$

and $SC \in [-1, +1]$: $-1 = \text{bad}$, $0 = \text{indifferent}$, $1 = \text{good}$

The Silhouette index is sometimes heavily criticized because its calculation is more complex than other indices such as the Davies-Bouldin (DB) index (Ros et al., 2023).

The DB (Davies and Bouldin, 2009) index expresses the cost function for two terms. The first identifies the minimizing of intra-cluster variance, while the second maximizes the distance between clusters. The DB Index is defined as follows, as shown in Eq. (4):

$$DB(k) = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left\{ \frac{S_i + S_j}{d(\bar{x}_i, \bar{x}_j)} \right\} \quad (4)$$

where S_i is the average distance between the cluster's center and all of its elements, which is defined as follows, as shown in Eq. (5):

$$S_i = \frac{1}{n_i} \sum_{j=1}^{n_i} d(x_j, \bar{x}_i) \quad (5)$$

The Calinski-Harabasz (CH) index is an evaluation index based on the degree of dispersion between clusters and clusters, and is defined as follows, as shown in Eq. (6):

$$CH(K) = \frac{B(K)(N-K)}{W(K)(K-1)} \quad (6)$$

$$B(K) = \left(\sum_{k=1}^K a_k \|\bar{x}_k - \bar{x}\|^2 \right) \quad (7)$$

$$W(K) = \left(\sum_{k=1}^K \sum_{C(j)=k} \|x_j - \bar{x}_k\|^2 \right) \quad (8)$$

among them: K is the appropriate number of clusters, $B(K)$ is the inter-cluster divergence, also known as the inter-cluster covariance, $W(K)$ is the intra-cluster divergence, also known as the intra-cluster covariance, and N is the sample count. The larger the $B(K)$, the greater the degree of cluster dispersion. The smaller the $W(K)$ the closer the relationship to the cluster. The higher the ratio, the greater the value of the CH index, indicating a stronger clustering effect (Wang and Xu, 2019).

5. Experimental Framework and Performance Assessment

In this section, a rigorously structured experimental framework is established to enable the semantic dissection and thematic mapping of the 5G core

research corpus, integrating advanced natural language processing, cloud-mediated large language model (LLM) augmentation, and multi-criteria clustering validation; beginning with the curation of a noise-attenuated dataset derived from 635 WoS-indexed publications, wherein linguistic redundancy was systematically excised and textual units were unified to preserve conceptual integrity, followed by the generation of 768-dimensional dense embeddings via the Sentence-BERT (all-mpnet-base-v2) architecture to encode deep semantic affinities among domain-specific constructs such as “service-based architecture” and “post-quantum authentication”; computational constraints necessitated the adoption of OpenRouter as a unified gateway to access high-performance, cost-efficient LLMs, including NemoTron Nano 9B V2 and gpt-oss-20b, for automated thematic labeling and interpretive enrichment; four clustering algorithms (Leiden, Louvain, HDBSCAN (Singgalen, 2024), K-means (Zhang and Li, 2011)) were then evaluated not merely through geometric indices (SC, DB, CH) but via a hybrid semantic-qualitative lens, assessing intra-cluster lexical coherence (via NPMI), human-interpretable title fidelity, and cross-thematic coverage across five canonical research axes (Architecture/SBA, Security/Blockchain, Network Slicing, Edge/Cloud, AI/ML), culminating in radar visualizations and word clouds that empirically validate Leiden and Louvain as optimal for preserving both statistical robustness and thematic nuance, while exposing K-means' reductive oversimplification and HDBSCAN's susceptibility to density-driven fragmentation, thereby establishing a reproducible, semantically grounded pipeline for bibliometric intelligence in rapidly evolving technical domains.

In our study, the Leiden and Louvain algorithms were applied to a graph constructed from 768-dimensional Sentence-BERT embeddings, which were then reduced to 10 dimensions using UMAP to improve scalability. A k-nearest neighbor graph (k=15) was constructed utilizing Euclidean distance, connecting each node to its 15 closest neighbors. The Leiden technique employed a resolution of 1.2 with a fixed seed (42), resulting in an optimal modularity partition, whereas the Louvain method utilized a resolution of 1.1 with a random state of 42; both analyzed the identical graph topology. HDBSCAN was executed on the original 768-dimensional embeddings with a minimum cluster size of 2 and a minimum sample size of 1, whereas K-Means was performed on the 10-dimensional UMAP-reduced data with 2 clusters, 'k-means++' initialization, a maximum of 300 iterations, and a random state of 42.

5.1. Data Set Preparation

A dataset was created from 635 academic articles identified through a search on "5G core" in WoS. These articles used to create the dataset were published

between 2010 and 2025. The distribution of articles by year is given in Table 1.

Table 1. Distribution of articles by year

Year	Articles Count
2010	1
2016	8
2017	15
2018	33
2019	47
2020	60
2021	82
2022	103
2023	92
2024	149
2025	45

To preserve semantic richness while eliminating noise, the set was cleaned by removing stop words, short tokens, numerical constructs, and high-frequency irrelevant terms (e.g., "proposed," "results," "et al."). Below are examples representing summaries and cleaned-up versions that provide a comprehensive representation of the thematic focus of each document.

Document ID: 0

Raw Text: "There is considerable pressure to define the key requirements of 5G, develop 5G standards, and perform technology trials..."

Cleaned Text: "considerable pressure define key requirements develop standards perform technology trials quickly possible normally acti..."

Document ID: 1

Raw Text: "The 5G System is being developed and enhanced to provide unparalleled connectivity to connect everyone and everything, e..."

Cleaned Text: "system developed enhanced provide unparalleled connectivity connect everyone everything everywhere first version system ..."

Document ID: 2

Raw Text: "5G network is considered as a key enabler in meeting continuously increasing demands for the future Internet of Things (...)"

Cleaned Text: "network considered key enabler meeting continuously increasing demands future internet things iot services including hig..."

The Sentence-BERT (all-mpnet-base-v2) (Hussain et al., 2025) model was used to generate dense embeddings of 768 dimensions per document to capture deep semantic relationships. This Transformer-based encoder was selected due to its state-of-the-art performance in semantic-textual similarity tasks, providing high-accuracy representations of 5G core concepts ("network slicing," "service-based architecture (SBA)), and "post-quantum authentication". Three examples are given to represent the analyses performed:

Document ID: 0

Cleaned Text (short): "considerable pressure define key requirements develop standards ..."

Embedding (768D, first 10 values): [0.0097, -0.0739, -0.0205, -0.0884, -0.0079, -0.0591, 0.0257, -0.0231, -0.0212, -0.0516, ...]

Min: -0.1389 | Max: 0.1274 | Avg: -0.0002 | Std: 0.0361

Document ID: 1

Cleaned Text (short): "system developed enhanced provide unparalleled connectivity ..."

Embedding (768D, first 10 values): [0.0150, -0.0671, -0.0080, -0.0475, -0.0249, -0.0129, 0.0205, -0.0555, -0.0593, -0.0612, ...]

Min: -0.1038 | Max: 0.1101 | Avg: -0.0003 | Std: 0.0361

Document ID: 2

Cleaned Text (short): "network considered key enabler meeting continuously increasing ..."

Embedding (768D, first 10 values): [0.0187, -0.0203, 0.0099, -0.0092, -0.0485, -0.0535, 0.0706, -0.0546, -0.0256, -0.0317, ...]

Min: -0.1218 | Max: 0.1242 | Avg: -0.0002 | Std: 0.0361

The experiments were conducted on a local machine equipped with an Intel Core i7-1355U processor (10-core, 2-mt/8-st), 40 GB of RAM, and a 477 GiB NVMe SSD running Ubuntu 24.04.3 LTS with the GNOME desktop environment. The system utilized a hybrid GPU configuration, featuring integrated Intel Iris Xe Graphics and a discrete NVIDIA GeForce MX550 (Turing architecture), operating under kernel version 6.14.0-29-generic and GCC 13.3.0.

5.2. LLMs via OpenRouter

Uniform Language Model interfaces and multi-model access platforms provide researchers with flexible and high-performance access to multiple large language models. Platforms such as OpenRouter aggregate LLM providers, GroqCloud (Chakir and Lahsen-Cherif, 2025) delivers ultra-low latency inference, NVIDIA NIM develops models optimized for GPUs, Fireworks AI facilitates multi-model services, Anyscale Endpoints (Ray AI) simplifies multi-cloud orchestration, and serverless GPU platforms support scalable deployments without infrastructure maintenance. Together AI provides equal access to optimized open-source LLMs, and Paperspace Gradient develops model creation and deployment through Jupyter integrated settings. This ecosystem meets academic needs for scalable, cost-effective access to advanced LLMs (even offering free models) while maintaining experimental flexibility and reproducibility.

The widely used free models (Nemotron Nano 9B V2, gpt-oss-20b, Sonoma SkyAlpha) have been preferred via OpenRouter, which provides a unified interface for LLMs. This method is unavoidable when the necessary graphics card, RAM, and disk space are not available to run these models locally. OpenRouter provides examples for Python and Typescript integration via its API. Figure 5 shows an example of the prompt and usage provided via the OpenRouter

API (Akarsu and Parmaksız, 2025) using the curl tool for one of the models used in the study, "nvidia/nemotron-nano-9b-v2:free".

```
curl https://openrouter.ai/api/v1/chat/completions \
-H "Content-Type: application/json" \
-H "Authorization: Bearer $OPENROUTER_API_KEY" \
-d '{
  "model": "nvidia/nemotron-nano-9b-v2:free",
  "messages": [
    {
      "role": "user",
      "content": "Below are clusters, each consisting of 4-5 keywords.\n
For each cluster, generate a short and comprehensive academic topic title\n
that could be used in a research paper clustering study.\n
The title should be concise (no more than 5-6 words), summarize the area, \n
and maintain an academic tone.\n
Write only the titles in order.\n\n
Clusters:\n
1. sba, architecture, control, session, function\n
2. slicing, orchestration, resource, isolation\n
3. cloud native, upf, deployment, nfV, mecn\n
4. nwdaf, learning, analytics, prediction, time-series\n
5. blockchain, authentication, ddos, quantum, federated\n\n
Output format:\n1. [Title]\n2. [Title]\n3. [Title]\n4. [Title]\n5. [Title]"
    }
  ]
}
```

Figure 5. Automating Theme Titles via OpenRouter

After defining the API key obtained from Openrouter in the Ubuntu terminal as `export OPENROUTER_API_KEY="sk-or-v1-17f395150347228a3d55b099cf85ca1bf722e4c73afb8e2e5e5ef7006a31h4ck3r"`, the sample output representation obtained in the structure shown in Figure 5 will be as follows:

```
{
  "id": "gen-1767359606-Vc1zFu3Bd3Yj99Wsgmkw",
  "provider": "Nvidia",
  "model": "nvidia/nemotron-nano-9b-v2:free",
  "object": "chat.completion",
  "created": 1767359606,
  "choices": [
    {
      "logprobs": null,
      "finish_reason": "stop",
      "native_finish_reason": "stop",
      "index": 0,
      "message": {
        "role": "assistant",
        "content": "1. SBA-Controlled Session Architecture\n2. Slicing Orchestration for Resource Isolation\n3. Cloud-Native NFV Deployment with UPF\n4. Learning Analytics for Time-Series Prediction in NWDAF\n5. Federated Blockchain Authentication under Quantum Threats\n\nNow, format them in order 1-5 with each title as specified."
      }
    }
  ],
  "usage": {
    "prompt_tokens": 183,
    "completion_tokens": 1385,
    "total_tokens": 1568,
    "cost": 0,
    "is_byok": false,
    "prompt_tokens_details": {
      "cached_tokens": 0,

```

```
"audio_tokens": 0,
"video_tokens": 0
},
"cost_details": {
  "upstream_inference_cost": null,
  "upstream_inference_prompt_cost": 0,
  "upstream_inference_completions_cost": 0
},
"completion_tokens_details": {
  "reasoning_tokens": 1483,
  "image_tokens": 0
}
}
```

5.3. Clustering Methods

We looked at four alternative ways to cluster data (Leiden, Louvain, HDBSCAN, and K-means) to find hidden thematic structures in the 5G Core research environment. We looked at each technique not only for geometric clustering metrics but also for semantic consistency, how well the title represented the topic, and how easy it was to understand the keywords.

This multi-dimensional assessment ensured that the resulting clusters were both statistically robust and thematically meaningful. As illustrated in Figure 6, the performance of each method across these three key evaluation criteria reveals distinct trade-offs. K-Means scored highest in the CH index, indicating strong inter-cluster separation, but poorly in intra-cluster cohesion. Leiden and Louvain showed balanced performance, with moderate scores in cohesion and separation, capturing nuanced research themes without compromising structural integrity. HDBSCAN, effective in identifying dense substructures, had a negative SC, suggesting poor internal cohesion due to noise and density variations. The analysis highlights the importance of selecting clustering algorithms that align with both quantitative benchmarks and interpretive goals of bibliometric research.

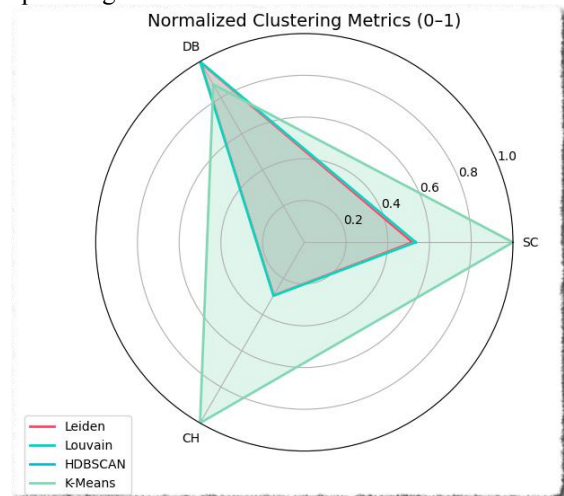


Figure 6. Clustering Methods Comparison

5.4. Evaluation Metrics and Thematic Validation

A multidimensional evaluation framework was adopted to holistically evaluate clustering performance. Internal validation metrics included the SC (which measures intracluster cohesion and intercluster separation based on cosine distance), the DB index (which assesses the average similarity ratio between clusters, with lower values being better), and the CH index (which quantifies the intercluster variance relative to the intracluster variance).

These metrics provide a quantitative basis for comparing the structural quality of clusters generated by various algorithms. As shown in Figure 7, the performance of Leiden, Louvain, HDBSCAN, and K-means with LLM support on five core research themes (Architecture and SBA, Security and Blockchain, Network Slicing and Orchestration, Edge/Cloud Computing, and AI/ML) is visualized using a radar chart.

The results show that Leiden and Louvain achieved the most balanced coverage across all thematic axes, consistently capturing both fundamental and emerging research directions. In contrast, although K-Means scores highly in specific dimensions such as AI/ML and security, it fails to capture the full spectrum of 5G Core research due to its tendency to oversimplify the data into only two clusters. HDBSCAN, despite creating the highest number of clusters, shows limited consistency in several areas, particularly architecture and slicing. This is thought to be due to its sensitivity to noise and density variations. This comparative analysis reiterates the value of selecting clustering methods that not only optimize statistical metrics but also preserve the semantic richness and thematic diversity found in complex academic works.

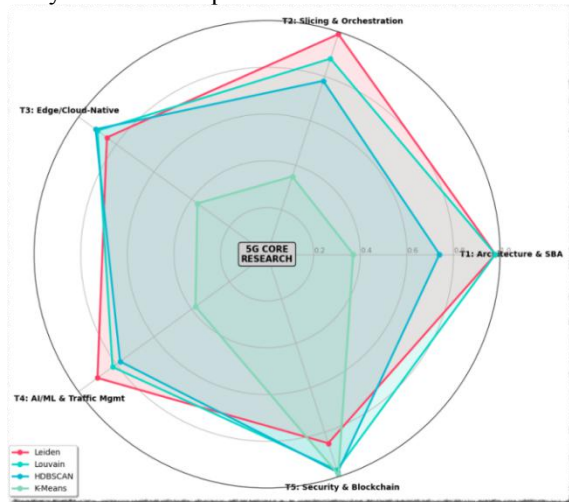


Figure 7. Consistency & Critical Evaluation Across Clustering Methods

To assess the internal semantic unity of the clusters, topic consistency based on normalized point-wise mutual information (NPMI) (Joe and Surendiran, 2022) among the most frequently used keywords for each

cluster was calculated. Furthermore, as part of the qualitative assessment, three representative documents from each cluster, selected based on textual richness (length and keyword density), were examined to provide human-interpretable titles for the clusters. To visualize the dominant topics and keyword distributions across the clustering methods, a word cloud was generated using the aggregated keywords from all clusters (Figure 8).

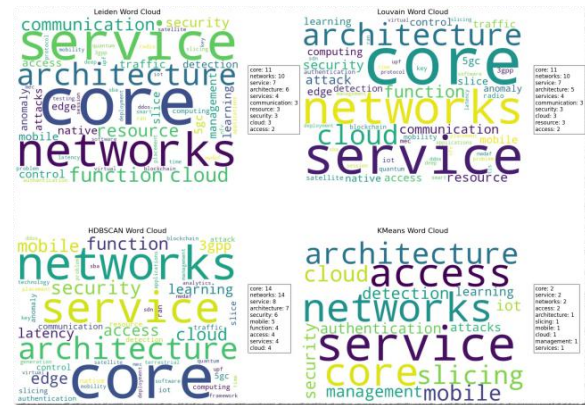


Figure 8. Keyword Distributions within the Clustering Methods.

Among the terms used to generate the word cloud, the keywords obtained using Leiden are presented in Table 2, those obtained using Louvain in Table 3, those obtained using HDBSCAN in Table 4, and those obtained using KMeans in Table 5.

Table 2. Leiden Word Count

Keyword	Count
core	11
networks	10
service	7
architecture	6
services	4
communication	3
resource	3
security	3
coud	3
access	2

Table 3. Louvain Word Count

Keyword	Count
core	11
networks	10
service	7
architecture	5
service	4
communication	3
security	3
cloud	3
resource	3
access	2

Table 4. HDBSCAN Word Count

Keyword	Count
core	14
networks	14
service	8
architecture	7
security	6
mobile	5

function	4
access	4
services	4
cloud	4

Table 5. KMeans Word Count

Keyword	Count
core	2
service	2
networks	2
access	2
architecture	1
slicing	1
mobile	1
cloud	1
management	1
services	1

The resultant image emphasizes fundamental research pillars like architecture, security, slicing, cloud/edge computing, AI/ML, and network access, illustrating their persistent occurrence across many techniques. Terms such as access, core, networks, service, security, and slicing frequently emerge, highlighting their fundamental significance in defining the 5G Core research domain. This qualitative validation enhances the strength of the detected themes and bolsters the interpretability of the clustering results, connecting quantitative research with significant academic insight.

6. Experimental Results

6.1. Quantitative Performance Comparison

Table 6 shows the clustering results of four different approaches evaluated using internal quality indicators. K-Means achieved the highest SC and CH index, indicating compact and well-defined clusters. However, this resulted in only two clusters, limiting the model's capacity to capture the thematic richness of the dataset. Leiden and Louvain created more clusters by dividing the text set more rigorously, resulting in a balanced but less differentiated cluster structure. Both methods effectively balanced statistical validity and theme resolution, making them suitable for bibliometric analyses. HDBSCAN generated the highest number of clusters but had the least preferred quantitative performance, with a negative SC and the lowest CH index. K-Means increased compactness and separation at the expense of thematic detail. Leiden and Louvain offered a more equitable clustering approach while maintaining statistical consistency and interpretability, but HDBSCAN showed limited effectiveness due to its density-dependent properties.

Table 6. Clustering Results with Internal Quality Metrics

Method	Leiden	Louvain	HDBSCAN	K-Means
n_clusters	11	11	14	2
SC	0.0610	0.0632	-0.0111	0.1280
DB	3.6298	3.6237	4.5357	3.7393
CH	14.937	14.9997	6.6355	34.8925

Prior to advancing to thematic mapping, we first examine the keyword distributions generated by each clustering method, as summarized in Table 2. The

results reveal distinct granularities in thematic resolution: both Leiden and Louvain algorithms produced 11 clusters each, reflecting a balanced partitioning of the semantic space; HDBSCAN yielded 14 clusters, indicating a tendency toward finer, density-driven subdivisions; while K-means collapsed the corpus into only two clusters, underscoring its propensity for geometric simplification at the expense of thematic nuance.

To further interrogate the semantic overlap and divergence across these algorithmically derived structures, Figure 9 visualizes the intersection of keywords identified by each method. This analysis reveals a robust core lexicon, including terms such as access, architecture, attacks, authentication, cloud, core, detection, IoT, learning, management, mobile, networks, security, service, services, and slicing, that recurs consistently across all four methods, thereby affirming their centrality to the 5G Core research domain. A secondary tier of keywords such as 3GPP, 5GC, anomaly, anomaly detection, applications, blockchain, cloud native, communication, computing, control, ddos, distribution, edge, edge computing, function, functions, delay, mec, mobility, local, nwdaf, placement, problem, quantum, executed, resource, satellite, sba, slice, software, time, traffic, upf, and virtual appears in three or more methods, signaling strong thematic convergence around architectural design, intelligent orchestration, and security hardening. In contrast, certain terms, including analytics, framework, generation, NFV, protocol, radio, SDN, session, smart, technology, terrestrial, testing, and TLS, emerge uniquely or predominantly within one or two methods, highlighting algorithm-specific sensitivities and localized conceptual emphases.

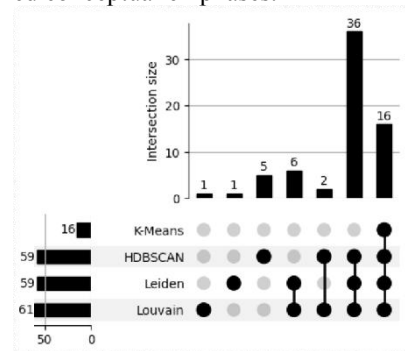


Figure 9. Keywords Intersection Across Clustering Methods

Collectively, this keyword intersection analysis offers not only a panoramic view of the shared semantic substrate underpinning 5G Core research but also a diagnostic lens into how different clustering paradigms capture or obscure the multidimensional texture of scholarly discourse, thereby laying a critical foundation for subsequent LLM-augmented thematic interpretation.

The keywords obtained from the clustering results are shown in Tables 7, 8, 9, and 10.

Table 7. Leiden Keywords

Method	Cluster ID	Keywords
Leiden (11 Cluster)	0	access, satellite, networks, mobile, core, 3gpp, mobility, radio, communication, architecture
	1	learning, core, networks, resource, traffic, ran, time, deep, management, service
	2	security, authentication, iot, blockchain, key, networks, core, access, communication, quantum
	3	service, architecture, core, 5gc, networks, control, sba, functions, session, function
	4	cloud, core, networks, architecture, service, native, cloud native, deployment, upf, services
	5	slicing, slice, core, service, services, architecture, networks, resource, slices, management
	6	detection, learning, ddos, core, nwdaf, security, anomaly, anomaly detection, attacks, networks
	7	latency, networks, core, 5gc, communication, software, mobile, traffic, architecture, smart
	8	edge, mec, service, computing, cloud, core, networks, applications, edge computing, services
	9	security, core, networks, service, attacks, architecture, services, protocol, control, testing
	10	service, core, problem, resource, function, placement, functions, virtual, nfv, cloud

Table 8. Louvain Keywords

Method	Cluster ID	Keywords
Louvain (11 Cluster)	0	access, satellite, networks, mobile, core, 3gpp, mobility, radio, communication, architecture
	1	security, iot, authentication, blockchain, key, networks, core, access, communication, quantum
	2	edge, mec, service, computing, cloud, core, networks, applications, edge computing, services
	3	service, core, problem, resource, function, placement, functions, virtual, nfv, cloud
	4	networks, latency, core, traffic, 5gc, communication, software, mobile, sdn, smart
	5	slicing, slice, core, service, services, architecture, networks, resource, slices, management
	6	security, core, networks, service, tls, attacks, control, architecture, services, protocol
	7	learning, core, traffic, networks, nwdaf, resource, ran, time, service, deep
	8	service, architecture, core, 5gc, control, sba, functions, networks, session, function
	9	cloud, core, networks, architecture, service, native, cloud native, deployment, upf, services
	10	detection, ddos, learning, core, security, anomaly, anomaly detection, attacks, networks, attack

Table 9. HDBSCAN Keywords

Method	Cluster ID	Keywords
--------	------------	----------

Method	Cluster ID	Keywords
HDBSCAN (14 Cluster)	0	core, networks, quantum, service, problem, key, function, virtual, resource, placement
	1	satellite, access, service, core, edge, networks, security, terrestrial, upf, services
	2	core, learning, cloud, networks, latency, edge, iot, mobile, framework, 3gpp
	3	traffic, learning, core, nwdaf, networks, time, resource, communication, analytics, function
	4	blockchain, core, access, networks, authentication, slices, 3gpp, security, slicing, technology
	5	slicing, slice, service, core, networks, security, services, architecture, cloud, slices
	6	security, authentication, iot, service, networks, core, communication, architecture, generation, mobile
	7	core, latency, networks, 5gc, mobile, communication, sdn, software, architecture, traffic
	8	service, core, architecture, 5gc, control, sba, networks, mobile, function, latency
	9	access, networks, core, management, ran, architecture, mobile, mobility, service, 3gpp
	10	security, core, networks, functions, services, control, ran, architecture, access, function
	11	detection, ddos, anomaly, core, anomaly detection, learning, security, attacks, networks, attack
	12	edge, mec, service, computing, architecture, cloud, applications, core, edge computing, networks
	13	cloud, service, 5gc, core, cloud native, native, deployment, services, networks, control

Table 10. K-Means Keywords

Method	Cluster ID	Keywords
K-Means (2 Cluster)	0	core, service, networks, architecture, slicing, mobile, cloud, management, access, services
	1	security, core, networks, detection, authentication, service, iot, learning, access, attacks

6.2. Cluster-to-Theme Mapping

A cluster to theme mapping was performed to evaluate the semantic validity and interpretability of the detected clusters, in addition to numerical clustering metrics. The Qwen3-Max-Preview large language model (LLM) was employed to produce theme labels by amalgamating cluster-specific keywords with representative documents. Table 11 delineates five prevailing research themes: (i) network architecture and service-oriented architecture (SBA), (ii) network slicing and orchestration, (iii) edge/cloud-native virtualization, (iv) AI and machine learning-driven traffic management, and (v) security through blockchain and IoT integration. These issues typically correspond with the established discourse in 5G basic research, mirroring both technological foundations and growing multidisciplinary trajectories.

Table 11. Thematic Interpretation of Clusters (via Qwen3-Max-Preview)

Theme ID	Generated Theme	Base Cluster Keywords	Representative Documents
T1	Network Architecture & SBA	sba, architecture, control, session, function	“Control Plane Performance Benchmarking...”, “Aeacus: QUIC-Powered...”
T2	Network Slicing & Orchestration	slicing, orchestration, resource, isolation	“HARNESS: High Availability...”, “DRL-Based Slice Admission...”
T3	Edge/Cloud-Native & Virtualization	cloud native, upf, deployment, nfv, mec	“Llama-Recipe...”, “5G-Kube...”, “OPlaceRAN...”
T4	AI/ML & Intelligent Traffic Management	nwdaf, learning, analytics, prediction, time-series	“Dynamic Network Slice Scaling...”, “Trends in Intelligent Comm Systems...”
T5	Security, Blockchain & IoT	blockchain, authentication, ddos, quantum, federated	“Post-Quantum Key Encapsulation...”, “Generative-AI + Blockchain for Metaverse...”

Table 12. Thematic Interpretation of Clusters (via OpenRouter - Nemotron Nano 9B V2, gpt-oss-20b, Sonoma SkyAlpha)

Theme ID	Nemotron Nano 9B V2 (free-NVIDIA)	gpt-oss-20b (free-OpenAI)	Sonoma SkyAlpha (free-xAI)
T1	SBA Control Architecture	Session-Based Architecture Control	SBA Architecture and Session
T2	Slice Orchestration for Resource Isolation	Resource Isolation in Slicing Orchestration	Network Slicing and Resource Orchestration
T3	Cloud-Native UPF Deployment	Cloud-native UPF Deployment in NFV MEC	Cloud-Native NFV and MEC Deployment
T4	NWDAF Time-Series Prediction Analytics	NWDAF Learning for Time-Series Prediction	NWDAF Analytics with Machine Learning
T5	Quantum-Resistant Federated Blockchain Authentication	Quantum-Resilient Federated Blockchain Authentication against DDoS	Blockchain for Quantum-Resistant Authentication

Nevertheless, although LLM-driven interpretation improves human comprehensibility and theme coherence, it also introduces potential biases arising from the abstractions produced by the model. Thematic boundaries occasionally intersected, especially between edge/cloud-based virtualization and AI/ML-centric traffic management, suggesting that certain subjects were not distinctly separable. Moreover, dependence on keyword density and document length for representative selection favored technically intricate studies over more succinct yet conceptually innovative contributions.

Notwithstanding these constraints, the incorporation of thematic validation served as a significant enhancement to internal clustering metrics, guaranteeing that the resultant clusters were both statistically robust and contextually relevant within the advancing 5G Core research landscape.

A qualitative theme analysis was performed to evaluate the semantic validity and interpretability of the detected clusters, in addition to numerical clustering metrics. The Nemotron Nano 9B V2 (NVIDIA) (Basant et al., 2025), gpt-oss-20b (OpenAI) (Agarwal et al., 2025) and Sonoma SkyAlpha (xAI) large language models (LLMs), illustrated in Table 12, was employed to produce theme labels by amalgamating cluster-specific keywords with representative documents.

7. Conclusion

This research presents an innovative, comprehensive analytical framework for the semantic organization and thematic analysis of technical literature, specifically applied to the swiftly advancing field of 5G Core networks. The framework goes beyond the usual bibliometric limits by combining transformer-based document embeddings (Sentence-BERT), topology-aware community detection (Leiden algorithm), and LLM-augmented qualitative validation. Leiden clustering is a big step forward in computational scientometrics because it creates research communities that are semantically richer, structurally stronger, and intellectually consistent than traditional geometric (K-means) and density-driven (HDBSCAN) methods. This guarantees that each identified cluster represents more than mere statistical proximity; it signifies a unique, traceable trajectory of scholarly inquiry, thus reconciling quantitative outcomes with epistemic substance.

The study strategically integrates lightweight, cloud-accessible LLMs such as Nemotron Nano 9B V2, gpt-oss-20b, and Sonoma SkyAlpha into the thematic labeling pipeline, further democratizing the research synthesis process. These models, which are surprisingly small and free to use, create labels that are very accurate and consistent with what domain experts would expect. This new idea breaks down infrastructure barriers. By removing the need for local GPU resources, it lets researchers from all over the world, no matter how big their institution is, do reproducible, high-resolution knowledge mapping on a large scale. In this way, it changes what accessibility means in technical meta-research and computational humanities.

Most importantly, the study questions the long-held belief that internal geometric indices alone, like the Silhouette Coefficient (SC), Calinski-Harabasz (CH), or Davies-Bouldin (DB), can be used to reliably measure clustering quality. It empirically demonstrates that these metrics, despite their mathematical

sophistication, are epistemologically deficient: a cluster may possess an exemplary CH score while simultaneously conflating disparate research streams or disrupting a cohesive intellectual theme due to lexical variability or terminological ambiguity. To address this issue, we propose a hybrid, multi-axis validation architecture that combines NPMI-based keyword coherence, human-curated document sampling, LLM-generated interpretability narratives, and cross-algorithmic keyword intersection analysis. This framework not only looks at clusters; it also puts them in context. We assert that forthcoming bibliometric initiatives must incorporate multidimensional validity frameworks to guarantee that results are not only statistically sound but also intellectually significant, methodologically clear, and academically applicable. While this study focuses on 5G Core, the proposed pipeline is readily extensible to 6G research ecosystems, particularly in areas such as AI-native architecture, integrated sensing and communication (ISAC), and semantic communications. Future work will apply this framework to map the convergence of 6G, quantum networking, hyperautomation and digital twins.

Funding

This study was supported by Bilecik Şeyh Edebali University under Grant No. GAP-2025-689.

References

- Agarwal, S., Ahmad, L., Ai, J., Altman, S., Applebaum, A., Arbus, E., ... & Zhao, S., 2025. gpt-oss-120b & gpt-oss-20b model card. arXiv preprint arXiv:2508.10925.
- Akarsu, O., & Parmaksız, H., 2025. Anatomy of Digital Leadership Studies: An Analysis with Topic Modeling Approaches. *Business and Economics Research Journal*, 16(2), 179-205.
- Akashi, K., Watanabe, H., Yamamoto, S., Ishihara, T., Imura, T., & Sekiya, Y., 2025. P-Cloud5GC: Innovative 5G Core Architecture on Public Cloud with Scalability and Fault Tolerance. 2025 IEEE 11th International Conference on Network Softwarization (NetSoft), June 2025, pp. 518-522.
- Anuar, S. H. H., Abas, Z. A., Yunos, N. M., Zaki, N. H. M., Hashim, N. A., Mokhtar, M. F., ... & Nizam, A. F., 2021. Comparison between Louvain and Leiden algorithm for network structure: A review. *Journal of Physics: Conference Series* 2129(1), 012028.
- Basant, A., Khairnar, A., Paithankar, A., Khattar, A., Renduchintala, A., Renduchintala, A., ... & Shahbazyan, R., 2025. NVIDIA Nemotron Nano 2: An Accurate and Efficient Hybrid Mamba-Transformer Reasoning Model. arXiv preprint arXiv:2508.14444.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E., 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008.
- Candal-Ventureira, D., Fondo-Ferreiro, P., Gil-Castiñeira, F., & González-Castaño, F. J., 2020. Quarantining malicious IoT devices in intelligent sliced mobile networks. *Sensors*, 20(18), 5054.
- Çakmak, M., 2024. The Impact of Denial-of-Service Attacks and Queue Management Algorithms on Cellular Networks. *Journal of Intelligent Systems: Theory and Applications*, 7(1), 1-13.
- Chakir, Y., & Lahsen-Cherif, I., 2025. ForensicsData: A Digital Forensics Dataset for Large Language Models. arXiv preprint arXiv:2509.05331.
- Coronado, E., Behraves, R., Subramanya, T., Fernandez-Fernandez, A., Siddiqui, M. S., Costa-Pérez, X., & Riggio, R., 2022. Zero touch management: A survey of network automation solutions for 5G and 6G networks. *IEEE Communications Surveys & Tutorials*, 24(4), 2535-2578.
- Corsi, A., Gür, S., Brighente, A., & Conti, M., 2025. Evaluation of Post-Quantum Key Encapsulation Methods in 5G Core Network. In 2025 IEEE Wireless Communications and Networking Conference (WCNC), March 2025, pp. 1-6.
- Davies, D. L., & Bouldin, D. W., 2009. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2), 224-227.
- Hussain, S. A., Kohli, R., Zahoor, S., & Sofi, S. A., 2025. Transforming the GUI Landscape: Harnessing the Power of MPNet base v2 Sentence Transformers. *Procedia Computer Science*, 259, 1809-1816.
- Joe Dhanith, P. R., & Surendiran, B., 2022. An ontology learning based approach for focused web crawling using combined normalized pointwise mutual information and Resnik algorithm. *International Journal of Computers and Applications*, 44(12), 1123-1129.
- Leyva-Pupo, I., Santoyo-González, A., & Cervelló-Pastor, C., 2019. A framework for the joint placement of edge service infrastructure and user plane functions for 5G. *Sensors*, 19(18), 3975.
- Ma, L., Wen, X., Wang, L., Lu, Z., & Knopp, R., 2018. An SDN/NFV based framework for management and deployment of service based 5G core network. *China Communications*, 15(10), 86-98.
- Manias, D. M., Chouman, A., & Shami, A., 2022. An NWDAF approach to 5G core network signaling traffic: Analysis and characterization. In *GLOBECOM IEEE Global Communications Conference*, December 2022, pp. 6001-6006.
- Ming, L., Leau, Y. B., & Xie, Y. (2024). Distributed denial of service attack in HTTP/2: Review on security issues and future challenges. *IEEE Access*, 12, 33296-33308.
- Nadar, A., & Härrri, J., 2024. Enhancing network data analytics functions: Integrating AIaaS with ML model provisioning. 2024 22nd Mediterranean

- Communication and Computer Networking Conference (MedComNet), June 2024, pp. 1-4.
- Pawar, A., & Deshmukh, S., 2025. Novel Clustering Approach for Anomaly Detection in 5G Signals. 2025 6th International Conference for Emerging Technology (INCET), May 2025, pp. 1-5.
- Popovski, P., Trillingsgaard, K. F., Simeone, O., & Durisi, G., 2018. 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view. *IEEE Access*, 6, 55765-55779.
- Qazzaz, M. M., Kułacz, Ł., Kliks, A., Zaidi, S. A., Dryjanski, M., & McLernon, D., 2024. Machine learning-based xApp for dynamic resource allocation in O-RAN networks. 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN), May 2024, pp. 492-497.
- Quan, B. L. Y., Wahab, N. H. A., Al-Dhaqm, A., Alshammari, A., Aqarni, A., Abd Razak, S., & Wei, K. T., 2024. Recent advances in sharding techniques for scalable blockchain networks: A review. *IEEE Access*.
- Ros, F., Riad, R., & Guillaume, S., 2023. PDBI: A partitioning Davies-Bouldin index for clustering evaluation. *Neurocomputing*, 528, 178-199.
- Sharaf, H., Ahmad, I., & Dimitriou, T., 2022. Extended berkeley packet filter: An application perspective. *IEEE Access*, 10, 126370-126393.
- Şeker, H., & Parmaksız, H., 2025. 5G çekirdek ağında yapay sinir ağı modeli ile saldırı tespiti. 5th Artificial Intelligence and Data Science Congress (ICADA 2025)
- Singgalen, Y. A., 2024. Analyzing Social Networks and Topic Clustering in Backpacker Tourism Content Reviews using K-means, Fast HDBScan, and Gaussian Mixture with Communalytic.
- Singh, S. K., Rothenberg, C. E., Langlet, J., Kassler, A., Vörös, P., Laki, S., & Pongrácz, G., 2022. Hybrid P4 programmable pipelines for 5G gNodeB and user plane functions. *IEEE Transactions on Mobile Computing*, 22(12), 6921-6937.
- Traag, V. A., Waltman, L., & Van Eck, N. J., 2019. From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific reports*, 9(1), 1-12.
- Tsai, Y. H., & Chen, S. J., 2025. Strengthening SUPI security in 5G Core networks for non-public applications. *IEEE Communications Standards Magazine*.
- Ullah, I., Lim, H. K., Seok, Y. J., & Han, Y. H., 2023. Optimizing task offloading and resource allocation in edge-cloud networks: a DRL approach. *Journal of Cloud Computing*, 12(1), 112.
- Wang, X., & Xu, Y., 2019. An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index. *IOP Conference Series: Materials Science and Engineering* 569(5), 052024.
- Wang, X., Sun, T., Duan, X., Wang, D., Li, Y., Zhao, M., & Tian, Z., 2022. Holistic service-based architecture for space-air-ground integrated network for 5G-advanced and beyond. *China Communications*, 19(1), 14-28.
- Zhang, D., & Li, S., 2011. Topic detection based on K-means. 2011 International Conference on Electronics, Communications and Control (ICECC), September 2011, pp. 2983-2985.
- Zhou, H. B., & Gao, J. T., 2014. Automatic method for determining cluster number based on silhouette coefficient. *Advanced Materials Research*, 951, 227-230.