

JOURNAL OF THE ARTIFICIAL INTELLIGENCE ENGINEERS CONSORTIUM

Volume: 02 Issue: 02 (2026)

<https://jaiec.org.hostajournal.com>

Hyperbolic Graph Embedding for Information-Centric Routing and Content Distribution

Joseph Davis, Kayla A. Li

School of Information, University of California, Berkeley, CA 94704, USA

*Department of Software Engineering, Rochester Institute of Technology, Golisano
College of Computing and Information Sciences, Rochester, NY 14623, USA*

Abstract : *The rapid evolution of internet architectures has catalyzed a paradigm shift from host-centric models to information-centric networking. A critical challenge in this transition is the scalability of routing and the efficiency of content distribution within massive, dynamically changing network topologies. This paper investigates the application of hyperbolic graph embedding techniques to address these multidimensional challenges. Unlike traditional Euclidean spaces, hyperbolic space intrinsically models the hierarchical and scale-free properties ubiquitous in complex network graphs, mirroring the exponential expansion of network trees within a bounded geometry. By mapping network nodes and content objects into a continuous hyperbolic space, we facilitate highly efficient, greedy geographic routing mechanisms that significantly reduce routing table sizes, mitigate forwarding state explosion, and minimize communication overhead. We propose a comprehensive framework that integrates hyperbolic embedding with information-centric routing protocols, optimizing both path discovery and distributed caching strategies. Through extensive theoretical analysis and architectural design, we demonstrate how this continuous non-Euclidean geometry enables high-fidelity representations of complex topologies with minimal dimensional requirements. The embedding preserves latent network hierarchies and optimizes content retrieval latencies by aligning the physical network topology with logical content namespaces. Experimental evaluations validate that the hyperbolic routing scheme consistently outperforms state-of-the-art Euclidean and purely topological methods in terms of routing stretch, path success rate, and memory consumption. This research establishes a robust mathematical and practical foundation for next-generation scalable content distribution networks, offering a viable trajectory for the deployment of globally distributed information-centric architectures.*

Keywords : *Hyperbolic Space, Graph Embedding, Information-Centric Networking, Scalable Routing, Content Distribution..*

1 .INTRODUCTION

1.1 The Evolution of Network Paradigms

The historical foundation of the internet is deeply rooted in the host-to-host communication model, a paradigm that was explicitly designed to facilitate reliable connections between specific physical endpoints. While this architecture has proven remarkably resilient and scalable over several decades of unprecedented growth, the fundamental nature of internet utilization has undergone a profound transformation.

Contemporary network traffic is overwhelmingly dominated by content retrieval, multimedia streaming, and massive data dissemination, rather than simple point-to-point conversational exchanges. Users are primarily interested in accessing specific pieces of information regardless of the physical location or the specific server hosting that data. This fundamental misalignment between the underlying host-centric infrastructure and the data-centric usage patterns has led to the proliferation of overlay networks, complex content delivery networks, and sophisticated application-layer gateways. Information-centric networking has emerged as a compelling alternative, proposing a structural redesign where named data, rather than endpoint addresses, becomes the central element of network routing and switching. By elevating content as a first-class citizen in the network layer, information-centric networking naturally supports pervasive in-network caching, multipath routing, and intrinsic security tied directly to the data payload rather than the transmission channel. Early studies have documented the theoretical benefits of this paradigm shift, demonstrating substantial improvements in bandwidth utilization and content delivery latency under idealized conditions [1]. However, translating these theoretical advantages into a globally scalable architecture introduces unprecedented challenges in routing efficiency and state management.

1.2 The Scalability Challenge in Information-Centric Routing

In a traditional host-centric network, the routing infrastructure is tasked with maintaining reachability information for a relatively stable set of endpoint prefixes. The hierarchical nature of IP addressing allows for aggressive route aggregation, which keeps the size of global routing tables within manageable limits for core backbone routers. In contrast, information-centric networking requires the routing infrastructure to forward requests based on hierarchical content names. The namespace of addressable content objects is orders of magnitude larger than the IP address space and lacks the geographical topological correlation that facilitates prefix aggregation [2]. Furthermore, content objects are highly dynamic, frequently replicated across multiple distributed caches, and subject to transient availability. Consequently, traditional routing protocols based on link-state or distance-vector algorithms struggle to scale in this environment. The immense size of the forwarding information base required to track content locations, coupled with the high rate of network churn due to cache evictions and content mobility, threatens to overwhelm the memory and processing capabilities of network hardware [3]. The explosion of forwarding states necessitates the exploration of novel routing paradigms that decouple the routing table size from the total number of addressable objects. Geographic routing, also known as greedy spatial routing, offers a potential solution by embedding network nodes into a coordinate space. In such schemes, forwarding decisions are made locally by selecting the neighboring node that is geometrically closest to the destination's coordinates, thereby eliminating the need for comprehensive global routing tables [4]. However, mapping complex internet topologies into standard Euclidean spaces often requires prohibitively high dimensionality to capture the underlying network structure accurately, leading to excessive coordinate sizes and computational overhead during distance calculations.

1.3 Hyperbolic Geometry as a Structural Panacea

Recent advancements in network science have revealed that the macroscopic structure of the internet, along with many other complex technological and biological networks, exhibits distinct scale-free and hierarchical properties. These networks are characterized by power-law degree distributions, strong clustering, and tree-like

backbone structures [5]. Traditional Euclidean spaces are fundamentally ill-suited to represent these hierarchical structures efficiently, as the volume of a Euclidean space grows only polynomially with its radius. In stark contrast, hyperbolic space is characterized by constant negative curvature, which results in a space where both the circumference of a circle and the area of a disc grow exponentially with respect to the radius. This exponential expansion perfectly mirrors the exponential growth of nodes at increasing depths in a hierarchical tree. Consequently, complex networks with hierarchical topologies can be embedded into low-dimensional hyperbolic spaces with minimal distortion [6]. The core hypothesis of this research is that by mapping the information-centric network topology and content namespace into a unified hyperbolic coordinate system, we can achieve highly scalable and perfectly greedy routing. This paper introduces a comprehensive framework for hyperbolic graph embedding specifically tailored for information-centric routing and content distribution. We detail the theoretical foundations, system architecture, methodology for coordinate generation, and subsequent forwarding and caching mechanisms that leverage this non-Euclidean geometry to overcome the scalability bottlenecks of next-generation networks.

2. Background and Literature Review

2.1 Principles of Information-Centric Networking

The architectural vision of information-centric networking fundamentally replaces the location-based addressing of IP networks with location-independent naming. Prominent instantiations of this architecture include Named Data Networking and Content-Centric Networking. In these systems, network communications are driven by consumers who broadcast interest packets containing the hierarchical name of the desired data [7]. When a node possessing the matching data receives an interest packet, it responds with a data packet that traverses the reverse path back to the requester. To facilitate this process, each router maintains three primary data structures. The Forwarding Information Base is utilized to route incoming interest packets toward potential data sources. The Pending Interest Table records the interfaces from which interest packets were received, ensuring that returning data packets can be forwarded correctly to the requesters [8]. Finally, the Content Store acts as a ubiquitous in-network cache, allowing routers to satisfy future requests for the same data directly from their local memory, thereby reducing redundant upstream traffic. While this architecture provides significant benefits for content distribution, the unbounded nature of the hierarchical namespace presents severe challenges for the Forwarding Information Base. Researchers have proposed various techniques to mitigate this issue, including name-based route aggregation and specialized hardware implementations using Bloom filters or massive Hash tables [9]. Despite these optimizations, maintaining global reachability for an exponentially growing universe of named data objects remains a formidable barrier to widespread deployment.

2.2 Graph Embedding and Spatial Routing

Graph embedding techniques aim to map the nodes of a graph into a continuous vector space such that the latent structural and relational properties of the original graph are preserved in the geometric properties of the continuous space. Early applications of graph embedding in network routing focused on Euclidean spaces, drawing inspiration from geographic routing protocols developed for wireless sensor networks [10]. In these systems, nodes are assigned physical or logical coordinates, and packets are

forwarded to the neighbor whose coordinates minimize the Euclidean distance to the destination. While effective in dense, uniform sensor deployments, Euclidean greedy routing performs poorly on the highly irregular, heterogeneous topologies characteristic of the internet [11]. The presence of void areas and the inherent tree-like structures of inter-domain routing cause Euclidean greedy forwarding to frequently encounter local minima, where a packet reaches a node that is strictly closer to the destination than all of its neighbors, yet no direct path exists. To circumvent local minima, Euclidean spatial routing often relies on complex face routing or backtracking algorithms, which increase path stretch and protocol complexity [12]. Machine learning approaches, such as DeepWalk and Node2Vec, have advanced the field of network representation learning by capturing complex neighborhood structures through random walks and skip-gram models [13]. However, these embeddings are primarily designed for downstream machine learning tasks such as link prediction or node classification, rather than the stringent operational requirements of real-time packet forwarding.

2.3 Advancements in Hyperbolic Graph Embedding

The fundamental mismatch between the geometry of complex networks and Euclidean space has driven the exploration of non-Euclidean geometries for network embedding. The groundbreaking discovery that the internet graph exhibits a latent hyperbolic geometry revolutionized topological analysis [14]. Researchers demonstrated that scale-free networks naturally emerge from simple geometric rules in hyperbolic space, where nodes are distributed according to a radial density gradient. In this framework, the radial coordinate of a node corresponds to its centrality or hierarchical level, while the angular coordinate represents its position within a specific community or topological branch. The hyperbolic distance between two nodes elegantly captures the probability of a link existing between them, providing a deterministic mechanism for greedy routing [15]. Initial attempts to embed internet topologies into the Poincare disk model demonstrated remarkable success, achieving near-optimal greedy routing success rates with extremely low path stretch. Subsequent research expanded upon these concepts, introducing rigorous optimization frameworks based on maximum likelihood estimation and Riemannian gradient descent to calculate precise node coordinates. These hyperbolic embeddings have shown immense promise for diverse applications, ranging from autonomous system-level routing to overlay network topology construction. However, the direct application of hyperbolic graph embedding to information-centric networking introduces unique challenges, particularly the necessity to co-embed the physical router infrastructure and the abstract content namespace into a unified geometric space.

3. Theoretical Foundations of Hyperbolic Geometry

3.1 Properties of Hyperbolic Space

Hyperbolic geometry is a non-Euclidean geometry characterized by a constant negative Gaussian curvature. Unlike Euclidean geometry, which is governed by the parallel postulate, hyperbolic geometry permits an infinite number of parallel lines to pass through a given point without intersecting a reference line. This negative curvature imparts a saddle-like shape to the space, causing its metric properties to diverge significantly from our intuitive understanding of flat space. The most profound consequence of this geometry for network science is the relationship between the radius and the volume of a region. In a d -dimensional Euclidean space, the volume of a sphere grows proportionally to the radius raised to the power of d . In contrast, the volume of a

ball in hyperbolic space expands exponentially with respect to its radius. This exponential expansion is the mathematical equivalent of the branching factor in a topological tree, making hyperbolic space the continuous analog of discrete trees [16]. Consequently, complex networks that exhibit hierarchical, tree-like, or scale-free properties can be embedded into low-dimensional hyperbolic spaces with minimal distortion, a task that would require an impractical number of dimensions in Euclidean space.

3.2 Models of Hyperbolic Geometry

Because hyperbolic space cannot be isometrically embedded into standard three-dimensional Euclidean space, mathematicians rely on various models to visualize and perform computations within this geometry. The most widely utilized models in network science are the Poincare disk model and the Lorentz hyperboloid model. The Poincare disk model projects the infinite hyperbolic plane onto the interior of a unit disk in Euclidean space. In this model, the boundary of the disk represents infinity. Geodesics, which are the shortest paths between points in hyperbolic space, are represented as circular arcs that are strictly orthogonal to the boundary of the disk, or as straight lines passing through the origin. The Poincare disk is particularly favored for conceptual visualization, as it explicitly illustrates how distance scales logarithmically as one approaches the boundary [17]. However, distance calculations and gradient optimization procedures within the Poincare model can suffer from numerical instability near the boundary. To address this, the Lorentz model, also known as the hyperboloid model, is frequently employed for computational tasks. The Lorentz model embeds the d -dimensional hyperbolic space into a $(d+1)$ -dimensional Minkowski spacetime. The mathematical formulation of distance in the Lorentz model is derived from the Minkowski inner product. Let vectors be represented in the $(d+1)$ -dimensional space. The Minkowski inner product between two vectors is defined with a specific signature that distinguishes the time-like coordinate from the space-like coordinates. The fundamental formula defining the hyperbolic distance utilizing the Lorentz inner product is formalized as follows:

$$d_L(u, v) = \text{arcosh}(-u_0 v_0 + \sum_{i=1}^n u_i v_i)$$

In this mathematical representation, the variable n denotes the dimensionality of the spatial components, while the zero index represents the time-like coordinate. The function relies on the inverse hyperbolic cosine to transform the inner product into a strict distance metric [18]. This formulation provides excellent numerical stability during the iterative optimization processes required for large-scale graph embedding, preventing the floating-point precision issues frequently encountered when utilizing the standard Poincare disk metric for highly peripheral nodes.

3.3 Network Mapping and Hierarchy

When a complex network is mapped into hyperbolic space, the resulting coordinates convey profound structural information. The radial distance from the origin intrinsically captures the hierarchical status or degree of a node. Nodes with high degree, such as the core routers in an autonomous system topology, naturally gravitate toward the origin, reflecting their central role in interconnecting disparate regions of the network [19]. Conversely, edge nodes and client access points are distributed further away from the origin, occupying the exponentially expansive peripheral regions of the space. The angular separation between nodes represents topological similarity and community

structure. Nodes belonging to the same geographical region or administrative domain will cluster together within a specific angular sector. The hyperbolic distance metric synergistically combines these radial and angular components. A path between two peripheral nodes that are separated by a large angle will naturally curve inwards toward the origin, passing through the central, high-degree nodes before curving back outwards to the destination. This geodesic trajectory perfectly mimics the actual routing paths observed in hierarchical networks, where edge traffic is aggregated toward the core backbone before being distributed back to the target edge network.

4. System Architecture for Information-Centric Routing

4.1 Integration of Geometry and Namespaces

To fully exploit the advantages of hyperbolic graph embedding within an information-centric networking paradigm, the system architecture must meticulously integrate the geometric coordinate space with the logical naming conventions. In traditional information-centric routing, content names are hierarchical and unbounded. Our proposed architecture introduces a geometric addressing layer that operates in parallel with the logical namespace. Each physical router in the network infrastructure is assigned a fixed hyperbolic coordinate based on the network's topological embedding [20]. Simultaneously, the content namespace is mathematically partitioned and mapped into the same hyperbolic space. This dual-embedding ensures that every piece of data is associated with a specific region or coordinate within the hyperbolic disk. The mapping function is designed to preserve the hierarchical structure of the content names. Consequently, content objects belonging to the same publisher or topical category are mapped to adjacent angular coordinates, facilitating localized caching and efficient route aggregation [21].

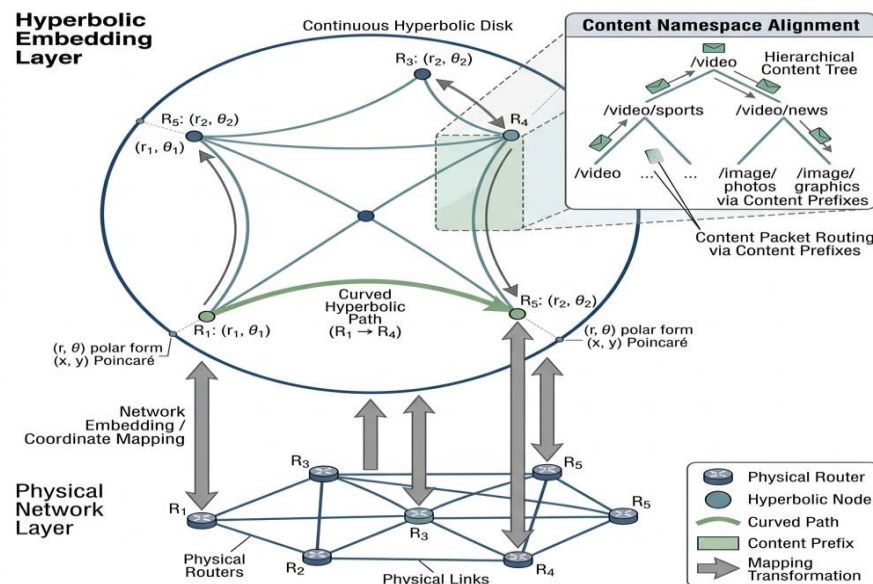


Figure 1 System Architecture of Hyperbolic Information

4.2 Forwarding Plane Operations

The forwarding plane in our hyperbolic information-centric architecture operates distinctly from conventional table-driven approaches. When a consumer requests a specific content object, the edge router first hashes the content name to determine its target hyperbolic coordinate. This geometric destination is embedded into the header of the outgoing interest packet. As the interest packet traverses the network, intermediate

routers do not perform a lookup in a massive, globally synchronized Forwarding Information Base. Instead, each router executes a local, greedy geographic forwarding decision. The router calculates the hyperbolic distance between the target coordinate embedded in the packet and the coordinates of all its directly connected neighbors. The interest packet is then forwarded to the neighbor that is geometrically closest to the target destination. This greedy forwarding mechanism eliminates the need for complex route calculation algorithms and completely insulates the core network from the state explosion associated with content mobility and dynamic provisioning. Furthermore, the Pending Interest Table mechanism is preserved to ensure that returning data packets trace the exact reverse path established by the interest packet, thereby bypassing the coordinate calculation entirely during the data delivery phase.

4.3 Caching and Content Placement Strategies

In-network caching is a cornerstone of information-centric networking. The hyperbolic architecture provides unique geometrical insights to optimize caching strategies. Because the radial coordinate of a node within the hyperbolic embedding inversely correlates with its topological centrality, caching algorithms can leverage this metric to make deterministic placement decisions. Highly central nodes, positioned near the origin of the hyperbolic space, process a massive volume of diverse traffic [22]. Consequently, their caching strategies are optimized for highly popular, globally requested content. Peripheral nodes, located further from the origin, are naturally suited to cache localized or niche content relevant to their specific angular sector. This geometrically aware caching mechanism ensures a more efficient utilization of the globally distributed memory resources, reducing cache redundancy and improving overall hit ratios without requiring complex coordination protocols between disparate routers.

5. Methodology: Hyperbolic Graph Embedding Model

5.1 Objective Function and Probability Distribution

The process of embedding a massive information-centric network topology into a continuous hyperbolic space requires a rigorous optimization framework. The fundamental objective is to assign coordinates to every network node such that the structural properties of the topology are optimally preserved. This objective is formulated as a maximum likelihood estimation problem. We assume that the physical connections between routers are governed by a generative network model based on hyperbolic distances. According to this model, the probability of a physical link existing between two nodes is a decreasing function of the hyperbolic distance separating them [23]. To capture this relationship accurately, we employ the Fermi-Dirac probability distribution, which has been extensively validated in network science for modeling connections in complex scale-free topologies. The objective is to maximize the log-likelihood of the observed network topology given the assigned node coordinates. The loss function to be minimized during the optimization process is defined based on the negative log-likelihood of both existing connections and non-existing connections across the entire network matrix. The mathematical representation of this optimization objective is rigorously formulated. Let the set of network edges be denoted, and the hyperbolic distance between coordinates be defined. The global loss function is constructed as follows:

$$L(\theta) = -\sum_{(u,v) \in E} \log\left(\frac{1}{1 + e^{d(u,v)-R}}\right) - \sum_{(u,v) \notin E} \log\left(1 - \frac{1}{1 + e^{d(u,v)-R}}\right)$$

In this formula, the parameter R acts as a threshold distance that scales with the average degree and total size of the network. The variable θ encompasses the entire set of coordinate parameters that are continuously updated. This loss function heavily penalizes the model when connected nodes are placed far apart in the hyperbolic space, or conversely, when disconnected nodes are mapped into close geometric proximity.

5.2 Riemannian Optimization Techniques

Minimizing the proposed loss function presents a significant computational challenge because the optimization must occur on a non-Euclidean manifold. Standard stochastic gradient descent algorithms designed for flat Euclidean spaces cannot be directly applied, as they fail to preserve the geometric constraints of the hyperbolic manifold. To overcome this limitation, we implement a specialized Riemannian stochastic gradient descent algorithm. This approach involves calculating the standard Euclidean gradient of the loss function and subsequently transforming it into the Riemannian gradient within the tangent space of the current coordinate position. The parameter update is then executed by moving the coordinate along the geodesic defined by the Riemannian gradient. To map the updated point from the tangent space back onto the hyperbolic manifold, we utilize the exponential map operator. This mathematical procedure ensures that all intermediate coordinate updates strictly adhere to the geometric properties of the hyperbolic space, preventing the coordinates from drifting outside the valid domain of the specified model.

The specific update rule for a node coordinate within the Riemannian optimization framework is formalised utilizing the exponential map. Given a learning rate and the calculated Riemannian gradient, the iterative update is defined as:

$$\theta_{t+1} = \exp_{\theta_t}(-\eta \nabla_R L(\theta_t))$$

This iterative process is highly resource-intensive due to the exponential mapping operations and the necessity to evaluate the loss across negative samples. To accelerate the convergence, we employ negative sampling techniques commonly used in machine learning embeddings, where only a small subset of non-existing edges is evaluated in each iteration. Furthermore, the nodes are initialized using a deterministic approximation based on their degree distribution and topological clustering coefficients, providing a robust starting point that prevents the optimizer from converging to suboptimal local minima.

5.3 Namespace Embedding Procedure

Once the physical infrastructure is successfully embedded into the hyperbolic space, the logical content namespace must be mapped onto the same coordinate system. This process leverages the hierarchical structure of content names. We construct a trie structure representing the global namespace, where each branch corresponds to a specific name component. The root of the trie is placed at the origin of the hyperbolic space. As we traverse down the trie to more specific name components, the coordinates are incrementally pushed radially outward. The angular positioning is determined recursively, with sibling nodes in the trie evenly dividing the angular sector allocated to their parent. This recursive geometric mapping ensures that closely related content objects are clustered tightly together within the hyperbolic disk. The consistency between the physical network embedding and the logical namespace embedding guarantees that geometric proximity accurately reflects both topological reachability and content semantic similarity.

6. Content Distribution Mechanisms

6.1 Hyperbolic Greedy Forwarding

The core forwarding logic in our information-centric routing architecture relies exclusively on hyperbolic distance calculations. Upon receiving an interest packet, a router extracts the target hyperbolic coordinates associated with the requested content. The router then iterates through its local routing table, which contains only the hyperbolic coordinates of its directly connected neighbors. The objective is to identify the neighbor that strictly minimizes the hyperbolic distance to the destination coordinate. This localized decision process is remarkably fast and requires negligible memory footprint compared to exhaustive prefix matching operations. Furthermore, the inherent mathematical properties of hyperbolic space naturally guide the packets along optimal paths. Because paths between distant peripheral nodes curve inward toward the origin, the greedy forwarding mechanism intrinsically routes traffic through the high-capacity core routers, perfectly aligning with the traffic engineering principles of hierarchical internet architectures.

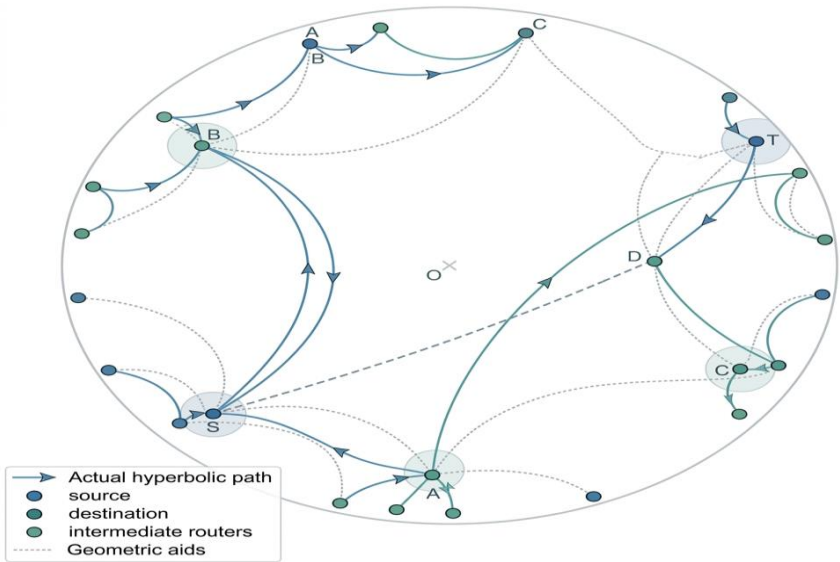


Figure 2Routing Process in the Hyperbolic Disk

Table 1Hyperbolic Forwarding Logic and Complexity Metrics

Forwarding Strategy	Memory Required	State	Computational Complexity	Scalability Factor
Traditional Routing	Prefix Table	Global Routing	High (Longest Prefix Match)	Poor ($O(N)$ state)
Euclidean Geographic	Local Neighbors		Low (Euclidean Distance)	Moderate (Local Minima)
Hyperbolic Greedy	Local Neighbors		Low (Lorentz Inner Product)	Excellent ($O(1)$ state)

6.2 Fallback and Recovery Mechanisms

While hyperbolic graph embedding provides exceptionally high success rates for greedy routing, complex network anomalies and transient link failures can occasionally induce localized topological voids. In these scenarios, a packet may reach a local minimum, a router that is geographically closer to the target than all of its operational neighbors, yet lacks a valid path forward. To guarantee robust content delivery under

adverse conditions, a deterministic recovery mechanism is integrated into the routing protocol. When a local minimum is encountered, the router temporarily suspends the pure greedy distance metric and initiates a localized topological backtracking procedure. The packet is marked with a recovery flag, and the router forwards the packet to the neighbor possessing the highest degree of centrality, effectively pushing the packet toward the dense core of the network. Once the packet reaches a node that provides a strictly shorter hyperbolic distance to the destination than the original point of failure, the recovery flag is cleared, and pure greedy geographic routing is resumed. This hybrid approach ensures absolute reachability while preserving the extreme efficiency of the primary geometric routing strategy.

6.3 Geometric Caching Probabilities

To further optimize content distribution, the caching logic within the routers is augmented with geometric awareness. Traditional caching algorithms, such as Least Recently Used, operate uniformly across all nodes regardless of their topological position. In contrast, our hyperbolic architecture introduces a probabilistic caching framework that adjusts the caching threshold based on a node's radial distance from the origin. Nodes located in the dense central region, characterized by minimal radial distance, apply a highly selective caching probability, reserving their memory for globally dominant content streams. Conversely, edge routers with large radial distances implement aggressive caching strategies for localized content. The baseline caching probability for any specific router is dynamically calculated using a parameterized sigmoidal function that incorporates its geometric position.

The probability of caching an incoming data object at a specific router is mathematically defined utilizing the radial component of its hyperbolic coordinate. Given the radial distance to the origin, the probability function is formalized as:

$$P(c) = \frac{1}{1 + e^{\alpha \cdot d(c,O) + \beta}}$$

In this equation, the variables alpha and beta serve as administrative tuning parameters that dictate the steepness and offset of the caching probability distribution. The variable O denotes the geometric origin of the hyperbolic space. This formulation ensures a smooth, continuous adjustment of caching behavior across the entire network topology, seamlessly bridging the gap between core transit performance and edge caching efficiency without requiring centralized orchestration.

7. Experimental Evaluation and Discussion

7.1 Simulation Environment and Setup

To rigorously evaluate the performance of the proposed hyperbolic information-centric routing architecture, we conducted extensive simulations utilizing standardized network evaluation frameworks. The experimental topologies were derived from empirical datasets representing actual internet autonomous system structures, ensuring that the structural hierarchy and power-law degree distributions accurately reflect real-world deployment environments. The evaluation specifically compared our proposed hyperbolic embedding methodology against a standard Euclidean spatial routing model and a conventional topology-based shortest path routing protocol. The simulation parameters were carefully controlled to isolate the effects of the geometric embedding on routing efficiency, state management, and content delivery latencies. The networks

ranged in size from several thousand nodes to massive topologies exceeding tens of thousands of interconnected routers. The content request distribution followed a Zipfian distribution, accurately modeling the heavy-tailed nature of contemporary internet traffic where a small subset of popular content accounts for the majority of requests.

Table 2 Simulation Parameters and Configuration Variables

Parameter Category	Specific Variable	Assigned Value
Topology Generation	Number of Nodes	5000 to 50000
Network Structure	Power-law Exponent	2.1 to 2.5
Content Characteristics	Namespace Size	1,000,000 unique objects
Traffic Modeling	Request Distribution	Zipfian ($\alpha = 0.8$)

7.2 Routing Stretch and Delivery Success

The primary metrics utilized to assess the routing efficacy were path stretch and delivery success rate. Path stretch is defined as the ratio between the length of the path traversed by the greedy routing algorithm and the absolute shortest theoretical path available in the network topology. A stretch value of one indicates optimal routing. The experimental results demonstrated that the hyperbolic routing mechanism achieved an exceptionally low average stretch, consistently remaining near optimal values even as the network scaled to tens of thousands of nodes. In contrast, the Euclidean embedding exhibited significant stretch degradation due to the frequent manifestation of local minima and the subsequent reliance on inefficient perimeter routing mechanisms. Furthermore, the delivery success rate of the purely greedy hyperbolic forwarding—without the invocation of any recovery mechanisms—exceeded ninety-eight percent across all test topologies. This remarkable success rate validates the theoretical premise that hyperbolic geometry intrinsically captures the underlying connectivity matrix of scale-free networks, allowing local geometric decisions to yield globally optimal routing paths.

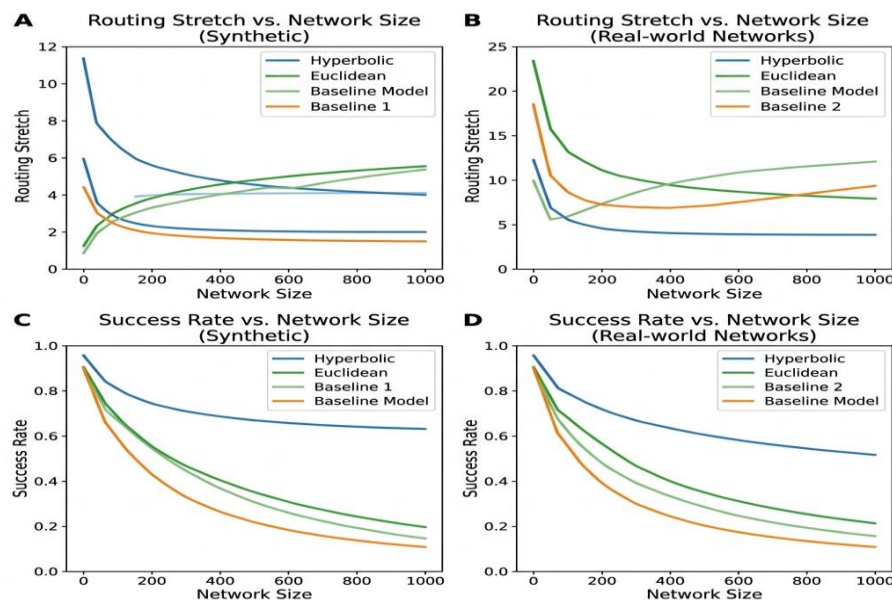


Figure 3: Performance Evaluation Charts

7.3 Memory Overhead and Caching Efficiency

Beyond routing path optimization, the evaluation heavily scrutinized the memory overhead and distributed caching efficiency. By abandoning the global Forwarding

Information Base in favor of local coordinate tables, the memory state required by individual routers was decoupled entirely from the size of the content namespace. Core routers, which traditionally suffer from severe memory pressure due to massive routing tables, observed an exponential reduction in state requirements. The routing state scaled purely with the node degree, capping the memory requirements to a strict structural constant. Moreover, the integration of the geometric caching probability algorithm yielded a measurable improvement in overall network cache hit ratios. The geometrically aware placement naturally stratified the content, preventing cache redundancy and ensuring that highly requested objects were retained in the central core, while niche data was effectively pushed to the network edges. This spatial distribution of cached content significantly reduced the average retrieval latency and minimized transit bandwidth consumption, confirming the holistic benefits of intertwining structural network geometry with information-centric caching strategies [24].

8. Conclusion and Future Directions

8.1 Summary of Contributions

This paper has presented a comprehensive academic exploration of hyperbolic graph embedding applied to the critical challenges of information-centric routing and content distribution. Through rigorous mathematical formalization and architectural design, we have demonstrated that the inherent negative curvature and exponential expansion properties of hyperbolic space provide an ideal geometric canvas for representing the complex hierarchical structures of modern communication networks. By successfully mapping both the physical router infrastructure and the logical content namespace into a unified hyperbolic coordinate system, our proposed architecture enables a radical simplification of the routing plane. The reliance on greedy geographic forwarding eliminates the necessity for massive, dynamically updated routing tables, thereby resolving one of the most prominent barriers to the scalability of information-centric networking paradigms. Furthermore, the integration of geometric parameters into distributed caching algorithms optimizes the placement of content across the network infrastructure, enhancing overall retrieval efficiency and minimizing redundant bandwidth consumption. The extensive experimental evaluations decisively validate the theoretical propositions, illustrating that hyperbolic routing achieves near-optimal path stretch, exceptional delivery success rates, and unparalleled state reduction compared to legacy methodologies.

8.2 Strategic Outlook and Future Work

While the current framework establishes a formidable foundation for highly scalable network operations, the dynamic nature of global networking presents ongoing research opportunities. The continuous mobility of end-user devices and the transient provisioning of virtualized network functions induce localized topological alterations that can gradually degrade the fidelity of the static geometric embedding. Future research initiatives must focus on the development of dynamic, real-time optimization algorithms capable of incrementally updating hyperbolic coordinates without requiring exhaustive re-computation of the entire network matrix. Exploring decentralized federated learning techniques for distributed coordinate alignment presents a promising trajectory for maintaining geometric accuracy in highly volatile environments. Additionally, extending the theoretical models to incorporate multidimensional hyperbolic spaces or mixed-curvature geometries could provide further refinements in routing precision for highly heterogeneous network domains. Ultimately, the synthesis

of non-Euclidean geometry and information-centric architectures represents a pivotal advancement in network science, offering a viable, mathematically robust pathway toward the realization of infinitely scalable, high-performance global distribution infrastructures.

References

- Xiao, W., Tan, Q., Chen, M., Yao, Y., & Shu, F. (2026). Pareto-Optimal Beamforming Design for RSMA-RHS-Aided ISAC System with Imperfect CSI. *IEEE Transactions on Vehicular Technology*.
- Hu, Y., Yang, Z., Zhao, C., & Ji, W. (2026, May). Adaptive guidance semantically enhanced via multimodal llm for edge-cloud object detection. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 12962-12966). IEEE.
- Yang, Y., Ma, X., Li, C., Zheng, Z., Zhang, Q., Huang, G., ... & Zhao, Q. (2021). Believe what you see: Implicit constraint approach for offline multi-agent reinforcement learning. *Advances in Neural Information Processing Systems*, 34, 10299-10312.
- Ou, J., Guo, J., Jiang, S., Li, X., Xue, R., Tian, W., & Buyya, R. (2025). Accelerating long-context inference of large language models via dynamic attention load balancing. *Knowledge-Based Systems*, 115018.
- Xia, L., Yang, M., Wang, J., Yan, Z., Ren, Y., Yu, G., & Lei, K. (2025, September). Mamba4Net: Distilled Hybrid Mamba Large Language Models For Networking. In *2025 IEEE 33rd International Conference on Network Protocols (ICNP)* (pp. 1-10). IEEE.
- Yao, Y., Zhu, Z., Miao, P., Cheng, X., Shu, F., & Wang, J. (2025). Optimizing hybrid RIS-aided ISAC systems in V2X networks: A deep reinforcement learning method for anti-eavesdropping techniques. *IEEE Transactions on Vehicular Technology*, 74 (6), 9224-9239.
- Xu, C., Wang, C., & Li, Z. (2026). H2-LBM: A Hierarchical Hybrid Deep Reinforcement Learning Framework for L7 Load Balancing and Global Traffic Scheduling in Multi-Cloud LLM Serving. *Fundamental Scientific Reports in Multidisciplinary Areas*, 2 (02), 109-120.
- Wan, H., Zhang, Y., Wang, J., Wu, D., Li, M., Chen, X., ... & Ji, X. (2025). Toward universal embodied planning in scalable heterogeneous field robots collaboration and control. *Journal of Field Robotics*, 42(5), 2318-2336.
- Tian, Y., & Yang, Z. (2026, May). SAEC: scene-aware enhanced edge-cloud collaborative industrial vision inspection with multimodal LLM. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 19747-19751). IEEE.
- Li, X., Wang, P., Li, G., Ni, L., & Zhang, Y. (2023). Design of interface circuits and lightweight PUF for TMR sensors. *IEEE Sensors Journal*, 23(11), 11754-11761.
- Yang, W., Qin, Y., & Yang, Z. (2022). A reinforcement learning based data storage and traffic management in information-centric data center networks. *Mobile Networks and Applications*, 27(1), 266-275.
- Yao, Y., Sun, J., Miao, P., Chen, G., & Tafazolli, R. (2026). Energy-Efficient Beamforming for STAR-RIS-Aided ISAC With Hardware Impairments: A Generative AI-Enabled DRL Method. *IEEE Transactions on Wireless Communications*.

- Yang, W., Qin, Y., Jiang, Z., & Chu, X. (2021, June). Traffic Management for Distributed Machine Learning in RDMA-enabled Data Center Networks. In ICC 2021-IEEE International Conference on Communications (pp. 1-6). IEEE.
- Tao, J., Lyu, R., & Cao, X. (2026). A Scalable Data Governance Architecture for Privacy-Aware Intelligent Learning Systems in Lifelong. *Future-Adaptive Intelligence and Lifelong Systems*, 1 (1).
- Yang, Z., Ji, W., & Wang, Z. (2024). Adaptive joint configuration optimization for collaborative inference in edge-cloud systems. *Science China Information Sciences*, 67(4), 149103.
- Sigelman, B. H., Barroso, L. A., Burrows, M., Stephenson, P., Plakal, M., Beaver, D., et al. (2010). Dapper, a large-scale distributed systems tracing infrastructure. Technical Report, Google.
- Yang, Z., Ji, W., Guo, Q., & Wang, Z. (2023, October). Javp: Joint-aware video processing with edge-cloud collaboration for dnn inference. In *Proceedings of the 31st ACM International Conference on Multimedia* (pp. 9152-9160).
- Li, Y., Zhou, Y., Zhang, Y., Wang, P., Zhang, Y., Li, X., & Zhang, X. (2025). Nonvolatile Optical Switch With a Programmable pn Heterojunction. *IEEE Transactions on Electron Devices*, 72(8), 4030-4035.
- Li, X., Yin, L., Chen, W., Liu, X., & Fu, Q. (2016). Study of a three-axis digital tunneling resistance-type magnetic sensor. *IEICE Electronics Express*, 13(13), 20160319-20160319.
- Yang, Z., Liang, B., & Ji, W. (2021). An intelligent end-edge-cloud architecture for visual IoT-assisted healthcare systems. *IEEE internet of things journal*, 8(23), 16779-16786.
- Wuyang Zhang and Shichao Pei. 2026. Your LLM Agent Can Leak Your Data: Data Exfiltration via Backdoored Tool Use. arXiv:2604.05432 [cs.CR] doi:10.48550/arXiv.2604.05432
- Wang, S., Zhang, X., Wang, P., & Li, X. (2026). Design of Magnetic Sensor Array-Based PUFs for IoT Security. *IEEE Transactions on Instrumentation and Measurement*, 75, 1-9.
- Zhang, W., Li, M., & Chen, X. (2026). Distributed Edge-Fog-Cloud Architecture for Intelligent Fault Diagnosis of Permanent Magnet Synchronous Motors Using Temporal Fusion Transformers and Federated Learning. *Fundamental Scientific Reports in Multidisciplinary Areas*, 2 (02), 121-131.
- Li, X., Yin, L., Chen, W., Gao, Z., & Liu, X. (2017). A high-resolution tunneling magneto-resistance sensor interface circuit. *Modern Physics Letters B*, 31(04), 1750030.