Optimal Hub Location Detection in Geo-distributed Networks

[1] Mrs Farhina S Sayyad, [2] Mrs Rushali S Navale, [3] Mrs Mrunali Patil,
[4] Mrs Jyoti Shivaji Pawar, [5] Mr Ashok Kumar Kalal, [6] Mr Hiraman S Jadhav
[1] Department of Computer Engg, DYPCOE Akurdi Pune, India
[2] Department of Computer Engg, DYPCOE Pune, India
[3] Department of Information Technology DYPCOE Pune, India
[4] Department of IT, SCOE Pune, jspawar, India
[5] Department of Computer Engg, APCOER Parvati Pune, India
[6] Department of Information Technology, DYPCOE Pune, India

Corresponding Author: **Hiraman S Jadhav**

Abstract: A recent report on location-based data reveals the growing potential within the dynamic local-mobile ecosystem. It examines how mobile ads powered by location data are gaining momentum across the industry, how location-aware features are boosting user engagement in mobile apps, and how local data is enabling countless small and medium-sized businesses to engage meaningfully in the mobile marketplace. The report also breaks down the fundamental technologies that power these capabilities and discusses important privacy issues. Location-aware tools have become a key driver of user interaction within apps. Major platforms like Facebook, Google, Yelp, Instagram, Groupon, Twitter, and others have adopted advanced location-based functionalities. Moving past the limited appeal of traditional "check-in" features, these apps now emphasize innovative approaches such as location-triggered alerts and intelligent, context-driven services. The era of globally distributed systems and services, determining optimal hub locations within geo-distributed networks has become critical for minimizing latency, improving service efficiency, and reducing operational costs. This paper investigates methods for identifying optimal hub locations that serve as central points for data aggregation, communication, or computation in geographically dispersed networks. We explore a combination of graph-theoretic models, clustering algorithms, and optimization techniques to evaluate hub placement strategies under varying constraints such as bandwidth, latency, load balancing, and fault tolerance. The proposed framework incorporates real-world geospatial data and network traffic patterns to simulate and validate performance. Our results demonstrate that intelligently selected hub locations can significantly enhance network performance and resource utilization, offering valuable insights for infrastructure planning in cloud computing, content delivery networks (CDNs), edge computing, and large-scale IoT deployments.

Keywords: Geo-distributed Networks, Hub Location Optimization, Network Topology Latency Minimization, Graph Theory ,Clustering Algorithms, Resource Allocation, Edge Computing, Content Delivery Networks (CDNs), Load Balancing, Geospatial Analysis

Introduction

The rapid expansion of cloud services, content delivery networks (CDNs), edge computing, and globally distributed applications has driven a growing demand for efficient and scalable network infrastructures. In such geo-distributed networks, determining the optimal location of hubs—which serve as central points for data aggregation, routing, or computation—is critical to ensuring low latency, high availability, and cost-effective operation.

Optimal hub location detection involves identifying nodes within a distributed network that can act as centralized or semi-centralized coordinators for surrounding nodes or regions. These hubs can significantly reduce communication overhead, improve response times, balance workloads, and enhance fault tolerance across the system. Applications range from edge computing deployments and distributed data centers, to IoT infrastructures, logistics systems, and telecommunication backbones.

However, the challenge of hub placement is complex due to multiple factors: diverse and dynamic network topologies, variable traffic loads, geographic constraints, infrastructure costs, and stringent performance requirements. Traditional solutions often rely on graph-theoretical models, facility location algorithms, or heuristic approaches, but these may not scale well or adapt to real-time conditions in modern distributed environments.

This paper aims to explore and evaluate methodologies for optimal hub location detection in geo-distributed networks, considering both theoretical models and practical constraints. We propose a framework that integrates geospatial analysis, network metrics (e.g., latency, bandwidth, and node centrality), and optimization algorithms to determine hub locations that maximize overall network efficiency. Through simulations and real-world data evaluations, we demonstrate the potential performance gains and provide insights for infrastructure planners and system architects.

I. Overview of Location Based Services

Location-Based Services (LBS) refer to a broad category of computer applications that utilize geographic data to manage or enhance specific functionalities. As information services, LBS play a growing role in modern mobile computing, especially in social networking and entertainment contexts. Accessible via mobile networks, these services leverage the real-time location of smart phones, tablets, and other devices to deliver context-aware features. With the continued rise of the mobile device market, LBS have become increasingly significant.

LBS applications span across numerous domains including healthcare, indoor navigation, entertainment, workplace productivity, and personal use. Common examples include locating nearby ATMs, tracking packages or vehicles, and determining the location of friends or employees. In the realm of mobile commerce, LBS are used to deliver location-specific advertisements or promotional offers. Additional use cases include localized weather updates, augmented reality games, and more—highlighting the convergence between telecommunications, computing, and geographic information systems.

Locating methods refer to the techniques used to determine the geographical position of a device, object, or individual. These methods are fundamental to the functioning of Location-Based Services (LBS), enabling systems to deliver context-aware and personalized services based on user or object location. The accuracy, coverage, cost, and energy efficiency of each method vary depending on the technology and application context.

Global Positioning System (GPS)

- Description: GPS uses a constellation of satellites to triangulate the position of a device equipped with a GPS receiver.
- Accuracy: Typically 5–10 meters outdoors.
- Advantages: High accuracy in open environments, global availability.
- Limitations: Poor performance indoors or in dense urban areas ("urban canyons").

Assisted GPS (A-GPS)

- Description: Enhances standard GPS by using data from network servers to improve speed and accuracy, especially in weak-signal areas.
- Use Case: Faster location fixes, especially on smart phones.
- Advantage: Combines satellite data with cellular networks for quicker results.

Wi-Fi Positioning

- Description: Determines location based on the proximity of known Wi-Fi access points.
- Accuracy: 5–50 meters indoors.
- Advantages: Works well indoors where GPS fails.
- Limitations: Requires a maintained database of Wi-Fi hotspot locations.

Cell Tower Triangulation

- Description: Estimates the device location based on signal strength and timing from nearby cellular towers.
- Accuracy: 100–1000 meters, depending on tower density.
- Advantages: Works anywhere there is mobile coverage.
- Limitations: Low precision compared to GPS or Wi-Fi.

Bluetooth and BLE (Bluetooth Low Energy)

- Description: Used for very close-range location tracking, often in indoor environments via beacons.
- Accuracy: Within 1–10 meters.
- Use Case: Indoor navigation, proximity marketing, retail tracking.
- Limitations: Limited range and coverage.

RFID (Radio Frequency Identification)

- Description: Uses electromagnetic fields to detect and track RFID tags.
- Accuracy: Centimeter-level in controlled environments.
- Use Case: Asset tracking, inventory management, indoor logistics.
- Limitations: Requires dedicated infrastructure and tags.

IP-Based Geolocation

- Description: Infers location based on the IP address of a connected device.
- Accuracy: 10–100 km (city-level granularity).
- Use Case: Regional content delivery, analytics, fraud detection.
- Limitations: Inaccurate for precise tracking, susceptible to VPN masking.

Sensor Fusion

- Description: Combines data from accelerometers, gyroscopes, magnetometers, and other sensors to estimate movement and improve location accuracy.
- Use Case: Enhanced positioning in GPS-deprived areas (e.g., tunnels or subways).
- Advantages: Improves continuity of location tracking.
- Limitations: Requires calibration and consumes device power.

Overview Proposed System

The proposed system aims to identify and deploy optimal hub locations in a geo-distributed network to minimize communication latency, balance network load, and enhance overall performance. The system integrates geospatial data, network topology analysis, and optimization techniques to make intelligent, data-driven decisions for hub placement.

System Overview

The system is composed of the following key components:

- **Data Collection Module**: Gathers geolocation data, network traffic patterns, node availability, and bandwidth usage from distributed nodes.
- **Network Modeling Engine**: Constructs a weighted graph representation of the network, where nodes represent devices or locations and edges represent connectivity with associated metrics such as latency or bandwidth.
- **Candidate Hub Selection**; Applies clustering techniques (e.g., K-Means, DBSCAN, or hierarchical clustering) to group nodes based on geographic proximity and network behavior, identifying potential hub candidates.
- **Evaluation & Feedback**: Uses simulation or real-time network testing to assess the performance of selected hub locations. Performance metrics include latency, throughput, packet loss, and resilience. Results are fed back into the model for iterative improvement.

Key Features

- **Dynamic Adaptability**: The system supports real-time re-evaluation of hub locations in response to changing network conditions, such as node failures or traffic spikes.
- **Scalability**: Designed to operate efficiently across large-scale distributed systems with thousands of nodes.
- **Multi-objective Optimization**: Balances trade-offs among conflicting objectives (e.g., latency vs. cost) using a weighted or Pareto-based approach.
- **Privacy and Security Awareness**: Incorporates privacy-preserving mechanisms when processing sensitive geolocation or user data, ensuring compliance with data protection regulations.

Application Domains

- **Edge Computing Infrastructure**: For deploying edge nodes in latency-sensitive applications.
- **Cloud Data Center Placement**: To optimize global or regional server locations.
- **Telecommunication Networks**: For optimizing routing hubs or signal relay stations.
- **IoT and Smart City Deployments**: To reduce communication delays in sensor networks.

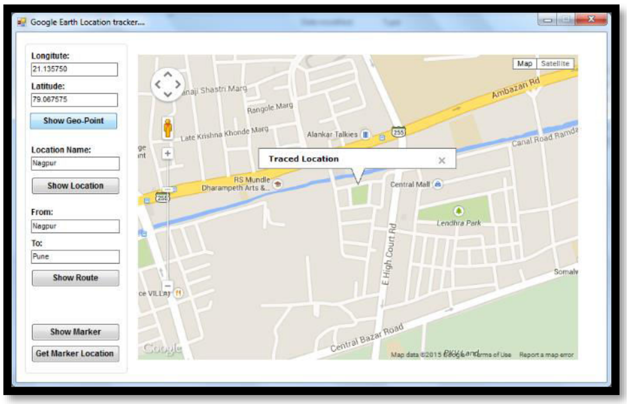

Fig 1.0 Expected Outcome & Future Scope

In recent years smart phone are become most important gadget for maintaining the daily activities and it also used by maximum population worldwide. Use of smart mapping technology is also increasing in large area like transportations, defense, sports, etc. Mapping applications are always depend upon current detection or preferred location of user or the group. Many application trying to get the user location to serve better service to location based services to user. Sharing location among group is better solution to know the individuals location. Finding or locating the location at known area or the known cities are usable and also feasible but at unknown location using these services may be risky or not feasible. Considering the condition any group wants to arrange a meeting at location which suits all the members hence it will always better to find centroid of the polygon generated by user geo-locations. It also has issue with finding better options of meeting while calculation. Proposed system aimed at finding the preferred and central location for user group using geo-point calculation and mapping technologies.

Figure 1.0 shows the overall working process of proposed system. This process includes multiple stages of execution. As per shown consider a condition there are five users in group planning to meet in centrally preferred location then one user from all will become master user and after which all user will share their location with master user and master user will execute the process. After execution system will calculate the central location by calculating the centroid of the polygon created by the user's connection. Once system get the central location it will ask user about his preferred location and after this using Google mapping API system will find out the nearest location selected by the user and once it found system will inform all user about final meeting location and if user wants he can view the travelling path to the location.

Methodology

Proposed system will employee following methodologies for successful completion of system

Great Circle Algorithm: The great-circle or orthodromic distance is the shortest distance between two points on the surface of a sphere, measured along the surface of the sphere (as opposed to a straight line through the sphere's interior). The distance between two points in Euclidean space is the length of a straight line between them, but on the sphere there are no straight lines. In non-Euclidean geometry, straight lines are replaced with geodesics. Geodesics on the sphere are the great circles (circles on the sphere whose centers coincide with the center of the sphere).

Through any two points on a sphere which are not directly opposite each other, there is a unique great circle. The two points separate the great circle into two arcs. The length of the shorter arc is the great-circle distance between the points. A great circle endowed with such a distance is the Riemannian circle.

Between two points which are directly opposite each other, called antipodal points, there are infinitely many great circles, but all great circle arcs between antipodal points have the same length, i.e. half the circumference of the circle, or πr, where r is the radius of the sphere.

The Earth is nearly spherical (see Earth radius) so great-circle distance formulas give the distance between points on the surface of the Earth (as the crow flies) correct to within 0.5% or so

Polygon Mid-Point Formulae

In geometry, the midpoint polygon of a polygon P is the polygon whose verrtices are the midpoints of the edges of P. It is sometimes called the Kasner polygon after Edward, who termed it the inscribed polygon "for brevity"

Triangle

The midpoint polygon of a triangle is called the medial triangle. It shares the same centroid and medians with triangle. The perimeter of the medial triangle equals the semiperimeter of the original triangle, and the area is one quarter of the area of the original triangle. The orthocenter of the medial triangle coincides with the circumcenter of the original triangle.

Quadrilateral

The midpoint polygon of a quadrilateral is a parallelogram called its Varignon parallelogram. If the quadrilateral is simple, the area of the parallelogram is one half the area of the original quadrilateral. The perimeter of the parallelogram equals the sum of the diagonals of the original quadrilateral.

Google Map API

After the success of reverse-engineered mashups such as chicagocrime.org and housingmaps.com, Google launched the Google Maps API in June 2005 to allow developers to integrate Google Maps into their websites. It is a free service, and currently does not contain ads, but Google states in their terms of use that they reserve the right to display ads in the future.

By using the Google Maps API, it is possible to embed Google Maps site into an external website, on to which site specific data can be overlaid. Although initially only a Java Script API, the Maps API was expanded to include an API for Adobe Flash applications (but this has been deprecated), a service for retrieving static map images, and web services for performing geocoding, generating driving directions, and obtaining elevation profiles. Over 1,000,000] web sites use the Google Maps API, making it the most heavily used web application development API.

The Google Maps API is free for commercial use, provided that the site on which it is being used is publicly accessible and does not charge for access, and is not generating more than 25 000 map accesses a day. Sites that do not meet these requirements can purchase the Google Maps API for Business.

The success of the Google Maps API has spawned a number of competing alternatives, including the Yahoo! Maps API, Bing Maps Platform, MapQuest Development Platform, and Open Layers.

Challenges and Future Outlook for Geo- Distributed Networks

Challenges

Despite the growing importance and applicability of optimal hub location detection in geo-distributed networks, several critical challenges must be addressed to ensure efficiency, scalability, and reliability:

- **Scalability and Complexity**: As the number of nodes in a distributed system increases, the complexity of hub location algorithms grows exponentially. Large-scale networks require computationally efficient algorithms capable of handling high-dimensional data and large geographic coverage areas without compromising accuracy.
- **Dynamic Network Topologies**: Many real-world networks, such as mobile, IoT, or vehicular networks, are highly dynamic. Nodes may frequently join, leave, or move within the network, making it difficult to maintain optimal hub locations over time.
- **Latency and Band width Constraints**: Optimizing for minimal latency must be balanced against limited bandwidth and the cost of data transmission, especially in remote or bandwidth-constrained areas. Ensuring consistent performance under varying network loads remains a challenge.
- **Multi-objective Trade-offs**: Hub placement often requires balancing multiple conflicting objectives—such as latency minimization, cost efficiency, fault tolerance, and energy consumption—making optimization a non-trivial task that may require heuristic or approximation methods.
- **Geographic and Environmental Constraints**: Physical and regulatory factors such as terrain, infrastructure availability, energy sources, and legal restrictions on data localization can impact the feasibility of deploying hubs in certain regions.
- **Data Privacy and Security**: The use of real-time location and network data raises significant concerns around privacy, especially when data involves users or sensitive systems. Secure, privacy-preserving computation methods must be integrated into hub location algorithms.

Future Outlook

The evolution of technologies like edge computing, 5G, AI-driven network management, and autonomous systems offers exciting opportunities for advancing optimal hub location detection:

- **AI and Machine Learning Integration**: Future systems may leverage AI/ML models to predict traffic patterns, user mobility, and network behavior, enabling more intelligent and adaptive hub placement decisions.
- **Real-time and Adaptive Optimization**: With the growth of dynamic, mission-critical applications (e.g., autonomous vehicles, smart cities), systems will need to support real-time hub reconfiguration to adapt to network changes without service disruption.

- **Decentralized and Collaborative Approaches**: Distributed algorithms and block chain-based coordination may allow nodes to collectively determine hub locations without relying on a central authority, enhancing resilience and scalability.
- **Sustainable and Energy-Aware Design**: Future hub location strategies will increasingly factor in energy efficiency and environmental sustainability, aligning with green computing goals.
- **Integration with Emerging Technologies**: The convergence of LBS, IoT, satellite-based internet (e.g., Starlink), and cloud-edge hybrid architectures will require more sophisticated hub location systems capable of operating across heterogeneous infrastructures.
- **Standardization and Interoperability**: Development of standardized protocols and frameworks for hub detection and coordination can promote interoperability among diverse platforms and vendors, accelerating adoption and innovation.

Recent Trends and Technological Innovations

The domain of geo-distributed network optimization and hub location detection has witnessed rapid advancements due to emerging technologies, increasing demands for low-latency services, and the growth of decentralized computing infrastructures. Below are some of the most notable trends and innovations shaping this field:

Edge and Fog Computing

The rise of edge and fog computing has shifted data processing closer to users, reducing the reliance on centralized cloud infrastructure. This paradigm necessitates strategic placement of edge hubs to support latency-sensitive applications such as autonomous vehicles, smart surveillance, and industrial automation.

- **Innovation**: AI-powered orchestration tools dynamically allocate workloads to optimal edge hubs based on real-time metrics.
- **Impact**: Improves response time, reduces backbone traffic, and enhances user experience.

AI and Machine Learning for Predictive Optimization

Artificial intelligence and machine learning algorithms are being increasingly applied to forecast user demand, mobility patterns, and network congestion. These insights support predictive hub location, where future network conditions are anticipated and pre-emptively optimized.

- **Innovation**: Reinforcement learning models continuously learn and adapt optimal hub placements based on feedback.
- **Impact**: Enhances adaptability and network resilience in dynamic environments.

Software-Defined Networking (SDN) and Network Function Virtualization (NFV)

SDN and **NFV** have enabled centralized control and dynamic configuration of network functions. These technologies allow virtual hub creation and reallocation without the need for physical infrastructure changes.

- **Innovation**: Virtual hubs can be spun up or migrated instantly based on network policies or user density.
- **Impact**: Increases flexibility and reduces costs for infrastructure deployment.

Integration with 5G Networks

The deployment of **5G** introduces ultra-low latency, high-speed connectivity, and network slicing capabilities. These enhancements enable highly granular hub optimization and support services with stringent real-time requirements.

- **Innovation**: Network slicing enables the creation of logical hubs optimized for specific services (e.g., IoT, streaming).
- **Impact**: Supports differentiated QoS and efficient resource utilization.

Real-Time Analytics and Telemetry

Advances in real-time monitoring and telemetry tools provide continuous feedback on network performance, user locations, and resource usage. This facilitates dynamic hub reconfiguration based on live data rather than static planning.

- **Innovation**: Use of streaming platforms like Apache Kafka and Apache Flink for processing network metrics in real time.
- **Impact**: Enables proactive decision-making and fault mitigation.

Block chain and Decentralized Coordination

Block chain technology is being explored for secure and decentralized hub coordination, especially in peer-to-peer and IoT networks. This removes reliance on centralized authorities and enhances trust among network participants.

- **Innovation**: Smart contracts automate hub selection and reward systems for cooperative nodes.
- **Impact**: Strengthens security, transparency, and auditability.

Sustainability and Green Networking

With increasing awareness of environmental impact, there is a shift toward energy-efficient hub location strategies. Systems are being designed to prioritize renewable energy sources, optimize power consumption, and minimize carbon footprint.

- **Innovation**: Energy-aware algorithms that consider power availability and consumption in hub placement decisions.
- **Impact**: Promotes environmentally responsible networking practices.

Multi-access Edge Computing (MEC)

MEC allows computation and storage at the edge of mobile networks, closely aligned with 5G. Hub location in MEC frameworks is critical for supporting low-latency mobile applications.

- **Innovation**: Telcos are integrating MEC hubs directly into cellular base stations.
- **Impact**: Drastically reduces latency and supports applications like AR/VR, telemedicine, and real-time gaming.

Conclusion

The increasing reliance on distributed computing systems, real-time applications, and location-based services has made optimal hub location detection a critical component of modern network design. As networks expand geographically and user demands evolve, the strategic placement of hubs becomes essential for minimizing latency, improving resource efficiency, enhancing user experience, and ensuring service reliability.

This paper has explored the fundamental principles, challenges, and technological advancements associated with hub location detection in geo-distributed networks. While traditional models offer a foundation, emerging technologies such as edge computing, AI-based optimization, 5G infrastructure, and software-defined networking are reshaping how hubs are selected, deployed, and managed. Despite promising progress, challenges related to scalability, dynamic topologies, multi-objective optimization, and privacy continue to persist.

References:

1. M. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman, 1979.
2. R. Jain, S. Paul, and J. Pan, “Network virtualization and software-defined networking for cloud computing: A survey,” IEEE Communications Magazine, vol. 51, no. 11, pp. 24–31, Nov. 2013.
3. J. G. Andrews et al., “What will 5G be?,” IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
4. F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, “Fog computing and its role in the internet of things,” in Proc. of the 1st Edition of the MCC Workshop on Mobile Cloud Computing, 2012, pp. 13–16.
5. Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing: The communication perspective,” IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2322–2358, 2017.
6. ISO/IEC 19762-5:2008, Information technology — Automatic identification and data capture (AIDC) techniques — Harmonized vocabulary — Part 5: Locating systems. ISO, 2008.

7. H. T. Kung and C. H. Wu, “Centrality and traffic flow in large-scale networks,” ACM SIGCOMM Computer Communication Review, vol. 36, no. 4, pp. 139–150, 2006.
8. D. Zhang, H. Xiong, L. Wang, and G. Chen, “Optimal location selection for hot new business in urban areas,” in Proceedings of the IEEE International Conference on Data Mining (ICDM), 2010, pp. 995–1000.
9. A. P. Rocha, T. R. de Oliveira, and E. F. Nakamura, “Energy-aware hub placement for wireless sensor networks,” Computer Networks, vol. 166, p. 106978, Jan. 2020.
10. L. Zeng, B. Zhang, X. Wang, and J. Wei, “Real-time hub placement in dynamic edge computing networks using deep reinforcement learning,” in Proceedings of IEEE INFOCOM 2021, pp. 1–10.
11. Igor Bilogrevic, Murtuza Jadliwala, Vishal Joneja,“ Privacy-Preserving Optimal Meeting Location Determination on Mobile Devices”, IEEE Transactions on Information Forensics and Security , Vol. 9, NO. 7,2014.
12. Rinku Dewri and Ramakrishna Thurimella, “Exploiting Service Similarity in Location Based Search Queries, “IEEE Transaction on Parallel and Distributed, February 2014.
13. Jing Liu, Zechao Li, Jinhui Tang, “Personalized Geo-Specific Tag Recommendation for Photos on Social Websites”, IEEE Transaction on Multimedia, Vol. 16, NO. 3, April 2014.
14. Linke Guo, Chi Zhang, “Privacy-Preserving Revocable Content Sharing in Geosocial Networks”, IEEE Conference on Communication and Network Security,2013.
15. Muhammad Ridhwan Ahmad Fuad and Micheal Drieberg, “Remote Vehicle Tracking System using GSM Modem and Google Map”, IEEE Conference on Sustainable Utilization and Development in Engineering and Technology, 2013.
16. Wei Li, Wei Jiao, Guangye Li, “A Location Privacy Preserving Algorithm for Mobile Lbs”, IEEE CCIS, 2012.
17. Wei Xin, Cong Tang, Tao Yang,Hui Ping sun, Zhong Chen, “Towards-Privacy Preserving RFID-Based Location Based Services”,IEEE International Conference on Fuzzy System and Knowledge Discovery, 2012.
18. Jianliang Xu, Xueyan Tang, “Privacy-Conscious Location-Based Queries in Mobile Environments”, IEEE Transactions on Parallel and Distributed Systems , VOL. 21, NO. 3, MARCH 2010 .
19. Hanunah Othman, Habibah Hashim, Jamalul-lail Ab Manan, “Privacy Preservation in Location-Based Services (LBS) Through Trusted Computing Technology”, IEEE International Conference on Communication, December 2009.
20. Leone C. Monticone, Richard E. Snow, “Minimizing Great-Circle Distance Ratios of Undesired and Desired Signal Paths on a Spherical Earth”, IEEE Transaction on vehicular technology, Vol. 58, NO. 9, November 2009.