CHAPTER 1
INTRODUCTION

1.1 Background

Blockchain Technology was Initially created as the basis for cryptocurrencies like Bitcoin, blockchain technology can now revolutionize a number of industries, particularly supply chain data sharing [1][2]. In the supply chain industry, timely access to data plays a critical role in ensuring efficiency and resilience of operations. Effective data sharing provides real-time informations into the supply chain information such as inventories, responsible person, and time. information obtained can be used by stakeholders for the benefit of the company.

However, conventional supply chain data systems often suffer from several limitations, including data inconsistency, siloed information, and inefficient query mechanisms when faced with large volumes of heterogeneous data. These limitations become more pronounced as organizations expand globally and the complexity of supply chain networks increases. As a result, stakeholders often struggle with delayed access to accurate data, leading to suboptimal inventory management and reduced responsiveness [3].

Blockchain-based platforms such as Hyperledger Fabric have been introduced to mitigate these issues by providing a decentralized, tamper-resistant infrastructure for data storage and sharing. Unlike public blockchain platforms, Hyperledger Fabric is designed for permissioned enterprise environments, where members are known and authenticated. It supports modular components such as pluggable consensus, chaincode (smart contract) execution, and configurable membership services, making it a flexible solution for a wide range of enterprise applications, including supply chain management [3].

Nevertheless, despite its modular and secure architecture, Hyperledger Fabric continues to face performance challenges in data retrieval, especially when operating over large-scale datasets. Querying the world state database—particularly in scenarios involving multi-condition filters or high data volume—can introduce substantial latency. The bottlenecks often occur during the endorsement and validation phases or as a result of inefficient query processing over unindexed fields [4][5].

Several studies have sought to address these performance limitations. For instance, Zhou et al. proposed LedgerData Refiner, an external data replication approach designed to improve query performance by mirroring blockchain data into off-chain databases [6]. Although effective in enhancing speed, this method introduces trust and consistency risks, as off-chain queries are no longer governed by the on-chain consensus and endorsement logic. On the other hand, Yan et al. presented an approach based on composite key design and LevelDB, demonstrating that structured keys can improve deterministic queries. However, their solution lacked support for rich query capabilities and was limited to simple key-value access patterns [7].

To overcome these limitations, this research proposes a hybrid query optimization strategy that combines Mango Query indexing and Composite Key construction within the CouchDB-based world state of Hyperledger Fabric. Mango Queries are well-suited for handling flexible, multi-field search conditions through JSON-based selectors and indexed fields. In contrast, composite keys enable high-speed, prefix-based lookups, particularly efficient for structured or hierarchical data access. By selectively applying the most suitable technique based on the complexity of the query—using Mango Queries for single-field filters and Composite Keys for multi-field lookups—this combined strategy seeks to enhance query performance while preserving data consistency and chaincode-level trust.

The proposed solution is evaluated using a simulated supply chain dataset of 10,000 and 50,000 records to reflect real-world blockchain data usage. By reducing query latency and improving scalability, the study aims to contribute to the development of more efficient data retrieval mechanisms for permissioned blockchain systems, enhancing their viability for large-scale industrial applications.

1.2 Theoretical Framework

This research is based on several main theories that underlie the development of query optimization strategies on Hyperledger Fabric blockchain systems. These theories include the concept of Distributed Ledger Technology, CouchDB database structure, Mango Query, and the use of Composite Key. The following is a description of the theories used.

1. Blockchain Theory and Distributed Technology (DLT)
   Blockchain is part of Distributed Ledger Technology (DLT), which allows recording transactions in a decentralized, transparent, and immutable manner. In the Hyperledger Fabric, DLT allows all licensed network members to participate in consensus and real-time tracking of asset [1]. This supports transparency and data integrity in supply chain management systems.

2. CouchDB and Mango Query on Hyperledger Fabric
   Hyperledger Fabric supports world state storage using CouchDB, a NoSQL document database that stores data in JSON format. Mango Query is a feature used to perform JSON selector-based searches, enabling flexible multi-attribute filters. Research by Zheng et al. (2022) showed that the performance of Mango Query is strongly influenced by the existence and structure of a suitable index [5].

3. Composite Key Theory in Chaincode
   Composite Key is a method of combining multiple attributes into one deterministic key to improve query efficiency. Hyperledger Fabric supports prefix-based search using the CreateCompositeKey function, which is proven to speed up data access when queries are performed on one specific attribute [3]. This approach is effective for simple search scenarios, such as by warehouse ID or product status.

1.3 Conceptual Framework/Paradigm

in this section there are 3 main points of the Conceptual Framework, such as Input, Process, and Output. Input is the data used for processing, which is in the form of inventory data containing *Product, Quantity, Warehouse, Status, Responsible person,* and *Timestamp*. In the process section, the data will be entered into the blockchain block using the private blockchain environment, Hyperledger Fabric. Next, is the output, the desired output is that when searching for data, the system will match the data with the Indices or Keys.

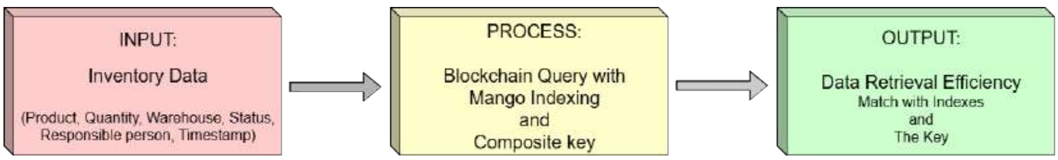

Fig. 1.1 Conceptual Framework.

1.4 Statement of the Problem

There are several problems with conventional Hyperledger Fabric query systems. This problem can certainly cause great disadvantages if it spreads to a larger scale. These problems are:

1. Low Data Retrieval Speed and Efficiency:
   In blockchain-based supply chain management systems such as Hyperledger Fabric, the efficiency of data retrieval is important because it directly affects the performance of the systems. Hyperledger Fabric has various query methods for data retrieval, including Mango Query and Composite Key. Each has advantages and limitations. Mango Query supports multi-attribute searches but is prone to performance degradation if the data volume is large or the index is not optimized, while Composite Key excels in single-attribute searches, but does not support the flexibility of complex queries. Reliance on any one method can lead to reduced performance, so an adaptive combination strategy is needed to overcome these limitations.

2. Limitations in data transparency:
   It cannot be denied that warehouse data is important because it includes data on goods, warehouses, and data on the person in charge of the warehouse. These data are very vulnerable to manipulation by irresponsible people.

1.5 Hypothesis

It is expected that the application of the combination strategy between Mango Query and Composite Key in the Hyperledger Fabric system can improve the efficiency of data retrieval compared to the use of each method separately. This strategy is expected to speed up query execution time, reduce resource consumption, and increase system scalability in supply chain management scenarios with large data volumes and various query complexities.

1.6 Assumption

1. The Hyperledger Fabric system has been properly configured and supports the use of CouchDB as the state database for Mango Query execution.

2. Mango indexes can be created and maintained regularly to support data search performance.

3. The key writing scheme for Composite Key has been designed with a consistent deterministic structure to support prefix-based search.

4. All experiments were run in a permissioned blockchain network environment with stable infrastructure specifications.

1.7 Scope and Delimitation

This research focuses on testing and evaluating the combination of Mango Query and Composite Key to improve data retrieval efficiency in the Hyperledger Fabric network. The scope of implementation is limited to the context of a supply chain management system, specifically related to inventory data that includes warehouse attributes, item status, person in charge, and recording time.

Experiments were conducted using one blockchain channel and one organization, as well as three organizations. The organization runs one peer instance connected to the CouchDB database as the state database. It should be noted that in the Hyperledger Fabric architecture, each peer has its own CouchDB, and as long as it is in the same channel and does not use private data, the contents of the world state in each CouchDB will be synchronous and identical.

This research also does not cover advanced security aspects such as private data collection, advanced encryption, or integration with external systems such as IoT. The main focus is on measuring query performance in terms of execution time, resource consumption, and efficiency of the data search strategy used.

1.8 Importance of the Study

This research aims to design, implement, and evaluate a combination strategy between Mango Query and Composite Key in data retrieval on Hyperledger Fabric permissioned blockchain network. This strategy is expected to improve data query efficiency in the context of supply chain systems, especially in scenarios with large data volumes and high complexity. The specific objectives of this research are:

1. Implement a data search function with a combination of Mango Query and Composite Key in a smart contract (chaincode).

2. Experimenting and measuring data retrieval performance based on time, CPU usage, and memory consumption.

3. Comparing the performance results of the combination strategy with conventional methods such as Mango-only and Composite-only.

1.9 Research Plan and Action Point

Month	Action Points
October 2024	Conduct initial research on Hyperledger Fabric, Mango Query, Composite Key, and blockchain-based inventory systems.
November 2024	Study the capabilities of CouchDB and Mango Query in Fabric; investigate prior optimization methods in rich queries.
December 2024	Set up Hyperledger Fabric testbed with CouchDB and implement chaincode for querying inventory data.
January 2025	Develop and test an adaptive query function combining Mango Query and Composite Key logic in Go chaincode.
February 2025	Design baseline functions; build workload scenarios for benchmark testing.
March 2025	Conduct single-user experiments; collect execution time, CPU usage, memory consumption, and result accuracy.
April 2025	Prepare experimental report and write research paper; analyze Mango vs Composite vs Adaptive performance.
May 2025	Extend experiment with Mango-only and Composite-only methods for deeper comparison and novelty enhancement.
June 2025	Finalize thesis document, including all results, analysis, discussion, and conclusions. Prepare for submission.