Hu-Fu: Efficient and Secure Spatial Queries over Data Federation

Yongxin Tong[1], Xuchen Pan[1], Yuxiang Zeng[2], Yexuan Shi[1], Chunbo Xue[1], Zimu Zhou[3], Xiaofei Zhang[4], Lei Chen[2], Yi Xu[1], Ke Xu[1], Weifeng Lv[1]

[1]State Key Laboratory of Software Development Environment, Beihang University, China,

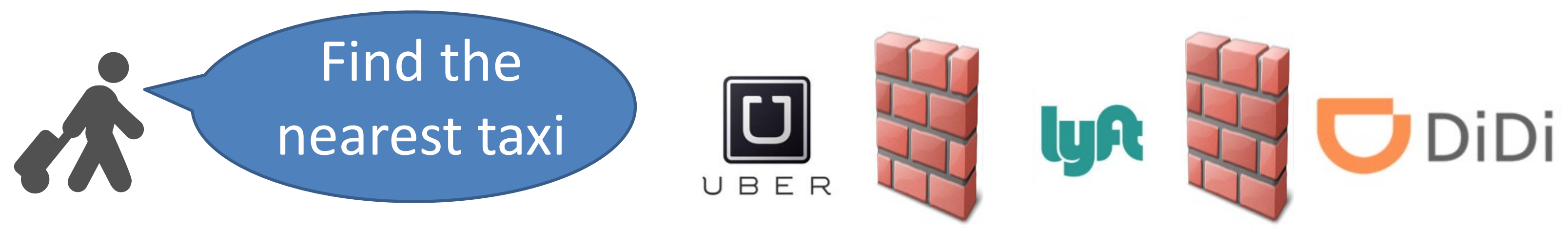
[2]The Hong Kong University of Science and Technology, [3]Singapore Management University, [4]University of Memphis

[1]{yxtong, panxuchen, skyxuan, xuechunbo, xuy, kexu, lwf}@buaa.edu.cn, [2]{yzengal, leichen}@cse.ust.hk,

[3]zimuzhou@smu.edu.sg, [4]xiaofei.zhang@memphis.edu

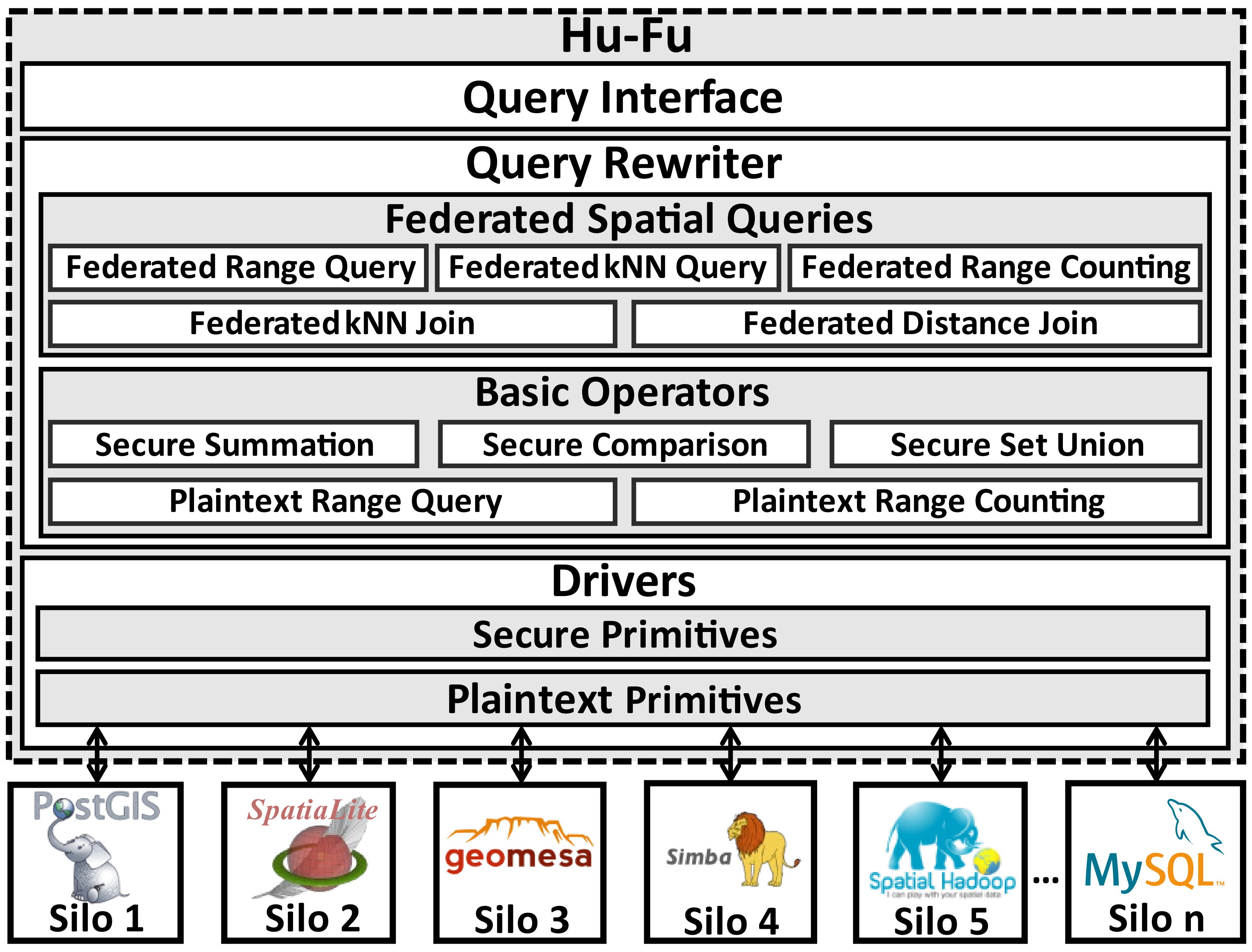
Introduction

- Spatial queries are essential for a wide spectrum of applications, but data isolation has become an obstacle to scale up query processing due to security concerns



- A promising paradigm to tackle the data isolation problem is to perform secure queries over a data federation
- Existing data federation systems are inefficient on spatial queries due to
  - excessive secure distance operations for query processing
  - usage of general-purpose SMC libraries for secure operation implementation

Hu-Fu Overview



Hu-Fu

Query Interface

Query Rewriter

Federated Spatial Queries

Federated Range Query | Federated kNN Query | Federated Range Counting

Federated kNN Join | Federated Distance Join

Basic Operators

Secure Summation | Secure Comparison | Secure Set Union

Plaintext Range Query | Plaintext Range Counting

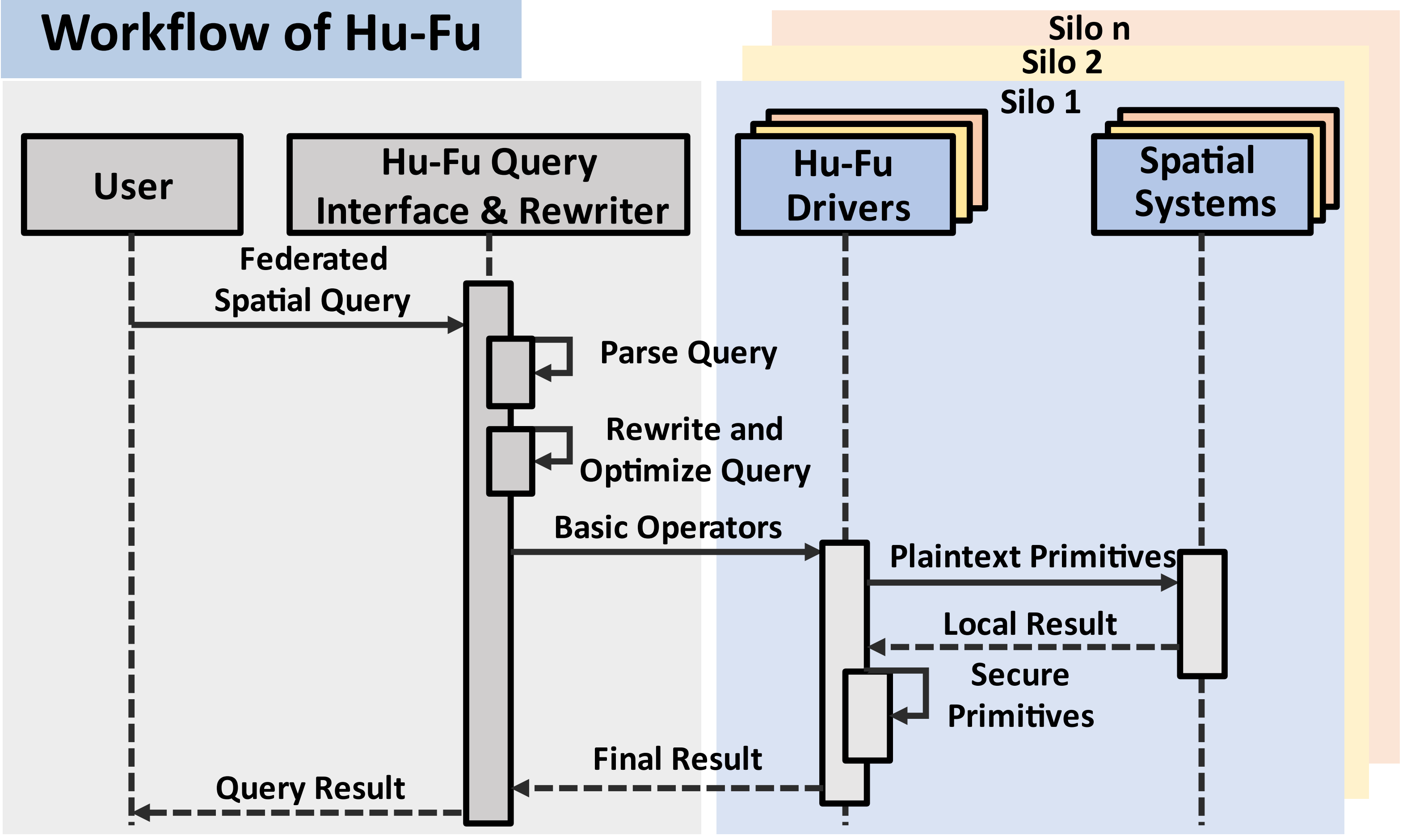
Drivers

Secure Primitives

Plaintext Primitives

- **Query Rewriter:** Decompose federated spatial queries into basic operators (plaintext operators and secure operators)
- **Drivers:** Implement secure operators as secure primitives with SMC protocol, and plaintext operators as plaintext primitives on top of silo's underlying spatial databases
- **Query Interface:** Provide federation view to users and support federated spatial queries written in SQL

Workflow of Hu-Fu



User

Hu-Fu Query Interface & Rewriter

Hu-Fu Drivers

Spatial Systems

Silo n

Silo 2

Silo 1

Federated Spatial Query

Parse Query

Rewrite and Optimize Query

Basic Operators

Plaintext Primitives

Local Result

Secure Primitives

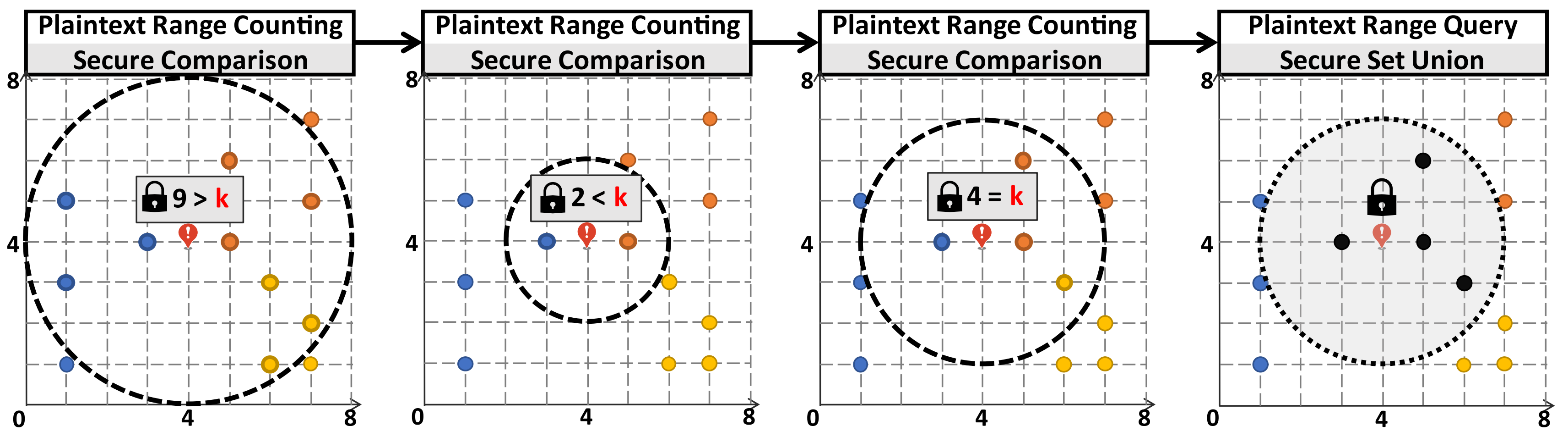
Final Result

Query Result

Query Decomposition

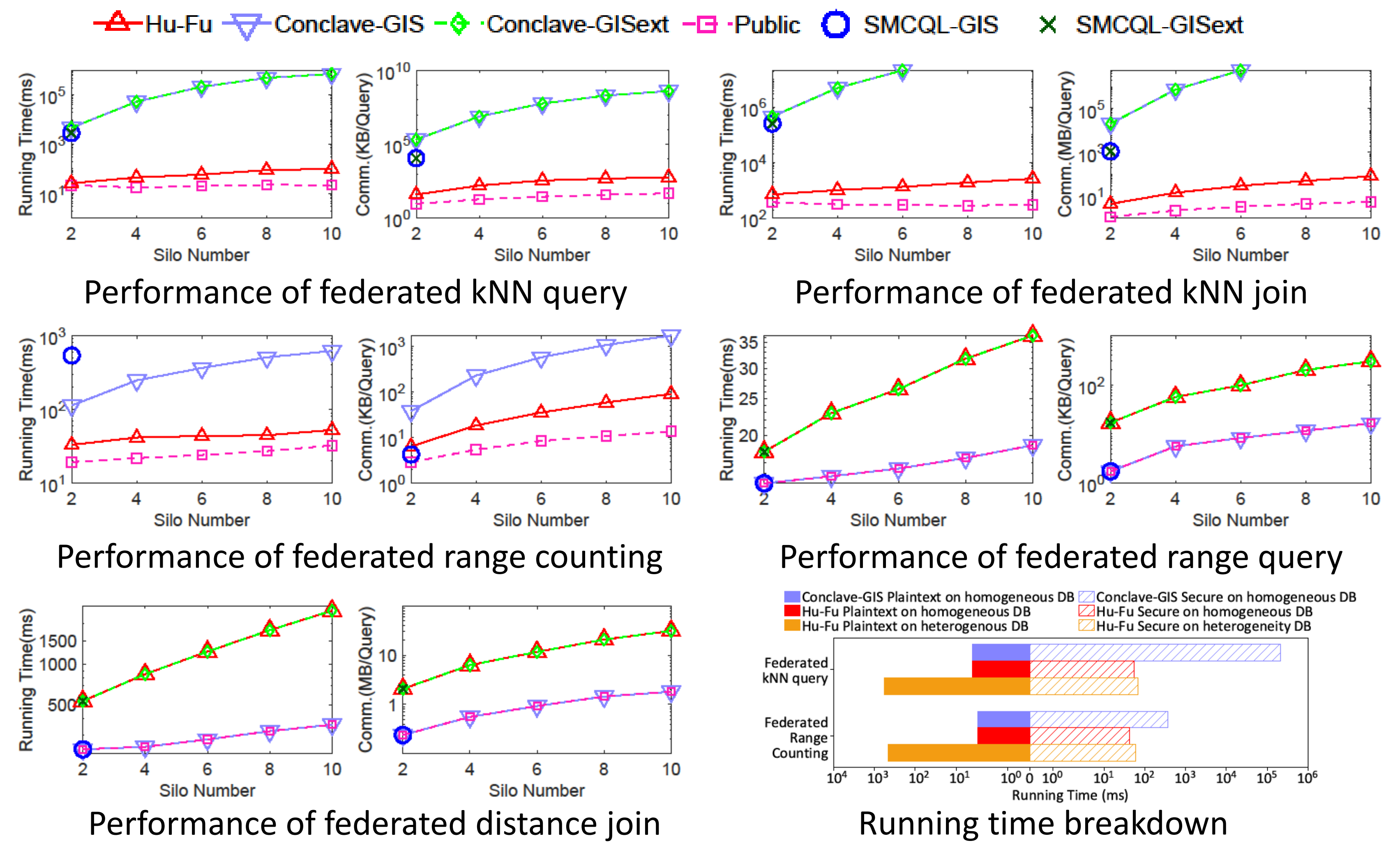
Category	Federated Spatial Query	Number of Plaintext Operator		Number of Secure Operator	
		Range Query	Range Counting	Comparison	Set Union / Summation
Radius-Known	Federated Range Query	n	0	0	$1/0$
	Federated Range Counting	0	n	0	$0/1$
	Federated Distance Join	$N\|R\|$	0	0	$1/0$
Radius-Unknown	Federated kNN Query	n	$O(n\, log\frac{v_0}{\epsilon_0})$	$O(log\frac{v_0}{\epsilon_0})$	$1/0$
	Federated kNN Join	$N\|R\|$	$O(\|R\|log\frac{v_0}{\epsilon_0})$	$O(\|R\|log\frac{v_0}{\epsilon_0})$	$1/0$

Decomposition principle: Decompose federated spatial queries into as many plaintext operators and as few secure operators as possible such that a large portion of the query can be executed in plaintext without compromising security



Example(Federated kNN query): We first derive a radius which contains k spatial objects via binary search, and then retrieve the spatial objects within this radius. In each binary search literation, we perform a plaintext range counting and a secure comparison to adjust the searching radius boundary. In the last round, a plaintext range query and a secure set union is performed to get the final result.

Experimental Evaluation



Performance of federated kNN query

Performance of federated kNN join

Performance of federated range counting

Performance of federated range query

Performance of federated distance join

Running time breakdown

Acknowledgment

We are grateful to anonymous reviewers for their constructive comments. Yongxin Tong's work is partially supported by the National Key Research and Development Program of China under Grant No. 2018AAA0101100, the National Science Foundation of China (NSFC) under Grant No. U21A20516, 61822201, U1811463 and 62076017, the State Key Laboratory of Software Development Environment Open Funding No. SKLSDE-2020ZX-07 and WeBank Scholars Program. This research was supported by the Lee Kong Chian Fellowship awarded to Zimu Zhou by Singapore Management University. Yuxiang Zeng and Lei Chen's work is partially supported by National Key Research and Development Program of China Grant No. 2018AAA0101100, the Hong Kong RGC GRF Project 16202218, CRF Project C6030-18G, C1031-18G, C5026-18G, AOE Project AoE/E-603/18, RIF Project R6020-19, Theme-based project TRS T41-603/20R, China NSFC No. 61729201, Guangdong Basic and Applied Basic Research Foundation 2019B151530001, Hong Kong ITC ITF grants ITS/044/18FX and ITS/470/18FX, Microsoft Research Asia Collaborative Research Grant, HKUST-NAVER/LINE AI Lab, Didi-HKUST joint research lab, HKUST-Webank joint research lab grants.