E-Commerce Letters 电子商务评论, 2026, 15(2), 971-980
Published Online February 2026 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2026.152238

Hans汉斯

面向电商风控的阈值触发隐私求交

秦颖琦，陈玉玲*

贵州大学公共大数据国家重点实验室，计算机科学与技术学院，贵州 贵阳

收稿日期：2026年1月16日；录用日期：2026年1月28日；发布日期：2026年2月28日

摘 要

面向电商跨平台风控联防与协同治理中交集稀疏但比对频繁的业务特征，传统多方隐私集合求交(MP-PSI)通常无条件执行完整流程，造成通信与在线时延开销冗余，并扩大不必要的隐私暴露面。为此，本文提出阈值门控多方隐私集合求交方案TG-MPSI，采用云辅助两阶段机制：第一阶段基于加密Bloom过滤器对交集基数进行安全估计，仅输出1比特触发标志；仅当估计规模超过预设阈值时，第二阶段才启动基于混淆布隆过滤器(GBF)与不经意传输(OT)的多方PSI计算以输出真实交集，用于联合拦截、黑名单协同与营销去重等任务。在半诚实模型下，方案满足输入隐私与结果正确性，并遵循最小披露原则。实验表明，在不同参与方数量、集合规模与阈值设置下，TG-MPSI能有效利用稀疏性在大量未触发场景显著降低总通信开销与在线延迟，相比无门控方案具有更好的工程可部署性与扩展性。

关键词

企业信息化，隐私计算，隐私集合求交，跨境电商，数据安全

Threshold-Triggered Private Set Intersection for E-Commerce Risk Control

Yingqi Qin, Yuling Chen*

State Key Laboratory of Public Big Data, School of Computer Science and Technology, Guizhou University, Guiyang Guizhou

Received: January 16, 2026; accepted: January 28, 2026; published: February 28, 2026

Abstract

To address the sparse-intersection yet frequent-matching pattern in cross-platform risk-control

*通讯作者。

文章引用: 秦颖琦, 陈玉玲. 面向电商风控的阈值触发隐私求交[J]. 电子商务评论, 2026, 15(2): 971-980.
DOI: 10.12677/ecl.2026.152238

alliances for e-commerce, conventional multi-party private set intersection (MP-PSI) protocols unconditionally run the full circuit, incurring redundant communication and on-line latency while unnecessarily enlarging the privacy-exposure surface. We propose TG-MPSI, a threshold-gated MP-PSI scheme that adopts a cloud-assisted two-phase architecture. In Phase 1, encrypted Bloom filters are used to securely estimate the intersection cardinality and return a single-bit trigger flag. Only when the estimated size exceeds a pre-defined threshold does Phase 2 activate, executing a full PSI based on garbled Bloom filters and oblivious transfer to output the real intersection for tasks such as joint blocking, blacklist collaboration and marketing deduplication. Under the semi-honest model, the protocol satisfies input privacy and result correctness while following the minimum-disclosure principle. Experiments show that, across varying numbers of participants, set sizes and threshold values, TG-MPSI exploits sparsity to cut total communication and online delay in the vast majority of non-triggered cases, offering better deployability and scalability than ungated baselines.

Keywords

Enterprise Informatization, Privacy calculation, Private Set Intersection, Cross-Border E-Commerce, Data Security

1. 引言

隐私集合求交(Private Set Intersection, PSI)源于密码学与安全多方计算(Secure Multi-party Computation, SMC)，旨在在不泄露各方输入集合的前提下安全计算交集。Meadows 等人基于 Diffie-Hellman 协议提出的两种方案首次系统性地引入了这一问题，为后续 PSI 研究奠定了基础[1]。随着数字经济与电子商务的发展，跨平台、跨部门、跨机构的数据协作需求日益突出：一方面，电商平台需在账户风控、黑产治理、刷单/羊毛党识别、商家准入与跨境合规等业务中对齐风险标识；另一方面，在网络营销、数据库营销及 *CRM* / *ERP* 联动等场景中也存在协同匹配需求。

也存在对“屏蔽名单/重复触达用户/异常流量设备”等标识进行交叉匹配的现实需求[2]。在这些场景下，直接共享用户标识、设备指纹或商家风险名单不仅涉及合规与隐私保护，还可能暴露商业敏感策略与运营数据，从而制约了跨域协作的深度与效率[1]-[4]。

早期 PSI 协议主要依赖公钥加密与 SMC 技术，通过加密、混淆以及不经意传输(Oblivious Transfer, OT)实现安全集合比较[3]；随后研究进一步整合同态加密与 Bloom 过滤器，在保障安全性的同时提升计算与通信效率[4]。近年来，PSI 也与区块链、可信执行环境(Trusted Execution Environments, TEEs)等技术结合，形成 2-PSI、MP-PSI 等多种变体，并在云端协同计算环境中得到更广泛讨论[5]-[7]。面向电商业务的大规模数据处理与快速响应需求，云辅助 PSI 为提升可扩展性提供了工程可行路径；但由于云环境可能不完全可信，云侧计算与中间结果仍可能带来隐私与策略泄露风险，使得安全的云端 PSI 成为关键挑战[8]。

在电子商务协同中，很多匹配任务并不需要“每次都输出精确交集”。例如，跨平台风控联动往往更关心“疑似重叠是否达到需要联动处置的规模”，数据库营销也常需要先判断“重复触达用户是否足以触发联合策略”。因此，为满足现实业务中“先评估价值、再决定是否深度协同”的需求，阈值 PSI (Threshold PSI)逐渐受到关注：仅当交集规模超过预设阈值时才执行后续的精确求交流程，从而避免低收

益协作带来的不必要开销[9]。然而，现有阈值 PSI 方案仍存在明显局限。多数方案在协议层仍近似采用“无条件执行”或高成本判定流程，即使在交集稀疏的电商场景下(例如平台用户群差异显著、风险标识较稀疏)也需承担较高的计算与通信开销[10]。这种强制执行不仅降低风控链路的实时性与系统吞吐，还可能引入时序、路径等侧信道泄露风险，违背“最小披露”原则。同时，在云辅助架构下，不可信基础设施带来的新型威胁进一步放大了上述问题。因此，如何在保证阈值判断与 PSI 执行安全性的前提下，实现更贴合电商业务的高效协同机制，是当前亟待解决的关键问题。

针对传统多方隐私集合求交(MP-PSI)在稀疏交集场景下仍需“无条件执行”而导致的高开销与隐私冗余问题，本文提出一种两阶段阈值门控多方 PSI 方案(Threshold-Gated Multi-Party PSI, TG-MPSI)，在不破坏 PSI 安全性的前提下实现“先判断是否值得计算，再决定是否执行”的自适应计算模式。本文主要贡献如下：

(1) 提出阈值触发的两阶段 PSI 框架，在多方 PSI 中引入可配置的交集基数阈值门控机制：当估计的交集规模未达到阈值时终止协议，从而避免电商跨平台协同中稀疏交集带来的无效计算与额外隐私暴露，实现“按需执行”的协同风控与联合治理模式。

(2) 设计基于同态加密与 Bloom 过滤器结合的安全阈值判定与条件 PSI 执行机制：云侧完成加密域聚合，领导者仅基于解密后的槽位统计进行基数估计，并结合 Chernoff 界实现阈值判定；在触发条件满足时，进一步结合零共享与混淆布隆过滤器(GBF)结构完成安全交集计算，适配电商场景下的云端可扩展部署。

(3) 在半诚实模型下给出基于仿真的安全性论证，保障输入隐私、输出正确性与阈值判定准确性。实验结果表明，在交集规模低于阈值时，TG-MPSI 相较无门控方案平均减少通信开销 30%~45%，在线延迟降低 25%~35%；当达到阈值时，整体性能与标准 PSI 同属一个量级，验证了方案在电商协同风控与数据协作中的可用性与扩展性。

本文其余部分组织如下：第 2 节概述相关工作；第 3 节介绍 TG-MPSI 的预备知识与问题定义；第 4 节描述系统模型与安全模型；第 5 节给出方案设计与正确性分析；第 6 节进行工程化落地；第 7 节总结全文并讨论后续工作。

2. 相关工作

现有 PSI 研究可概括为三条主线。

(1) 基于密码学原语的 PSI：利用 OT 扩展与同态加密实现大规模集合比较，代表性协议在通信与计算复杂度上持续优化[11]。这类方案安全表达能力强，但在多方云环境下的工程成本与运维复杂度仍需权衡。

(2) 基于数据结构的高效 PSI：Bloom 过滤器及其派生结构通过常数时间的插入与查询支持高效匹配。GBF/OBF 通过混淆、随机填充与路径隐藏等机制提升对抗侧信道的能力，并进一步扩展到并集/交集基数计算与多方场景[12] [13]。这类方法更贴近企业信息化中的高并发匹配需求，但需谨慎处理泄露边界与业务语义之间的平衡。

(3) 云辅助与阈值/门限 PSI：云计算为企业提供弹性算力，但不可信基础设施引入新的隐私风险与治理挑战[8] [14] [15]。阈值 PSI 将“是否执行”与阈值判断结合，适用于风控触发、合规核验与营销触达等业务逻辑，代表性方案结合 Paillier、阈值投票与零知识证明实现门限判定[16] [17]。但在企业常见的稀疏负载下，部分方案仍缺少稳定的“按需降本”效果，或因证明链路较重而增加系统集成成本。

针对电商稀疏匹配的真实负载特征，本文提出 TG-MPSI，以轻量的云端基数判定作为门控层，在不影响触发场景求交能力的前提下显著降低未触发场景的通信与在线计算，并提供企业信息化集成与可运

维部署路径。

3. 预备知识与符号

3.1. 符号约定

本文常用符号及含义如表 1 所示。

Table 1. Table of protocol notation
表 1. 符号表

符号	含义
P_i	第 i 个参与方
$P=P_{i}^{n}{}_{i=1}$	参与方集合， n 为参与方数量
C	云服务器(提供聚合与并行计算能力)
S_i	参与方 P_i 的私有集合(电商风险标识集)
λ	安全参数/字符串长度(bit)
k	哈希函数个数
$H=h_{j}{}_{j=1}^{k}$	k 个独立哈希函数集合
m	Bloom/GBF 的长度(槽位数)
T	触发阈值(门控条件)
$Enc(\cdot)/Dec(\cdot)$	同态加密/解密算法
$flag$	触发标志(是否进入第二阶段)
Sim	安全证明中的模拟器

3.2. Bloom 过滤器与混淆布隆过滤器(GBF)

Bloom 过滤器是一种空间高效的概率型数据结构，可用于快速判定元素 x 是否属于集合。其由长度为 m 的比特数组 $BF[1..m]$ 和 k 个相互独立的哈希函数 $H=h_{j}{}_{j=1}^{k}$ 组成，其中 $h_j:\mathcal{U}\to 1,\cdots,m$。初始化时 BF 全为 0。插入元素 x 时，对每个 $j\in[1,k]$ 置位 $BF[h_j(x)]=1$。查询 x 时，若存在某个 j 使得 $BF\left[h_j(x)\right]=0$，则 x 一定不在集合中；若所有对应位均为 1，则判定 x 可能在集合中(存在误报概率)，该误报概率通常随 m 、 k 以及集合规模变化，但在合理参数下可控制在较低水平。

混淆布隆过滤器(Garbled Bloom Filter, GBF)，GBF 是在 Bloom 过滤器基础上引入“混淆编码”的结构化 PSI 构件[18]。与仅存储比特不同，GBF 维护长度为 m 的字符串数组 $GBF[1..m]$，每个槽位存放 λ 比特串。对元素 x (可视为 λ 比特串)，令其索引集合为 $I_x=h_1(x),\cdots,h_k(x)$。编码时在 I_x 对应的 k 个槽位写入 k 份随机份额，使得这些份额按 XOR 还原为 x，即满足：

$$\bigoplus_{j\in I_x} GBF[j]=x \tag{1}$$

加法同态加密本文第一阶段的阈值判定采用加法同态加密以支持“加密域聚合、明文端判定”的计算模式。以 Paillier 方案为例，其满足对明文加法的同态性：对任意明文 m_1,m_2，有

$$Dec\left(Enc(m_1)\cdot Enc(m_2)\right)=m_1+m_2\ (\mathrm{mod}\ N) \tag{2}$$

并且对任意常数 a，

$$Dec\left(Enc\left(m_1\right)^a\right)=a\cdot m_1\left(\mathrm{mod}\,N\right) \tag{3}$$

其中， N 为 Paillier 公钥模数， $\cdot$ 与幂运算在密文群中进行。该性质使云端能够对密文进行加和聚合，而不直接接触各方输入，符合电商跨域协作中最小披露、可扩展计算的工程要求。

4. 模型描述

本节面向电子商务协作场景，给出阈值触发 TG-MPSI 的系统模型、输入输出、通信架构与阶段划分，并据此定义安全模型。在电商 C2C/B2B 业务中，多平台需对风险标识集合进行交叉核验以支撑联防联控、黑产治理、营销去重与商家准入，但直接共享名单易带来合规与商业敏感信息泄露风险。为此，TG-MPSI 采用两阶段门控机制：阶段一对交集规模进行隐私保护估计，未达阈值则终止；达到阈值时触发阶段二输出真实交集，从而在保障隐私的同时降低稀疏交集下的冗余开销。

系统由两类实体构成：

电商参与方：记为 $P=P_1,\cdots,P_n$ 。参与方 P_i 持有私有集合 $S_i=x_1^i,\cdots,x_{|S_i|}^i$ ，其中元素 x 为标准化编码的电商风险标识(账号 ID、设备指纹、商家 ID 等)。可执行本地哈希映射、GBF/Bloom 编码及加解密等操作。

云服务器：记为 C 。负责执行与阈值判定相关的结构化计算与密文域聚合。考虑到电商协作中云基础设施可能属于第三方服务商或多租户环境，本文将 C 建模为半诚实(semi-honest)：遵循协议但可能试图从所见消息中推断额外信息。

为便于工程实现，指定某一参与方 P_ℓ 为领导者(Leader)，负责收集来自云端的阈值判定结果、在触发条件满足时协调第二阶段 PSI 的执行并分发最终交集结果。

云辅助多方隐私集合求交协议：

输入： n 个参与方的私有集合 $S_1,\cdots,S_n$ 以及阈值 T 。

输出：触发标志 $flag$ 与条件输出结果 R ：

若 $flag=1$ ，输出真实交集 $R=\bigcap_{i=1}^{n}S_i$ ；

若 $flag=0$ ，输出 $R=\perp$ (表示不触发精确求交)。

传统 MPSI 系统模型如图 1 所示。

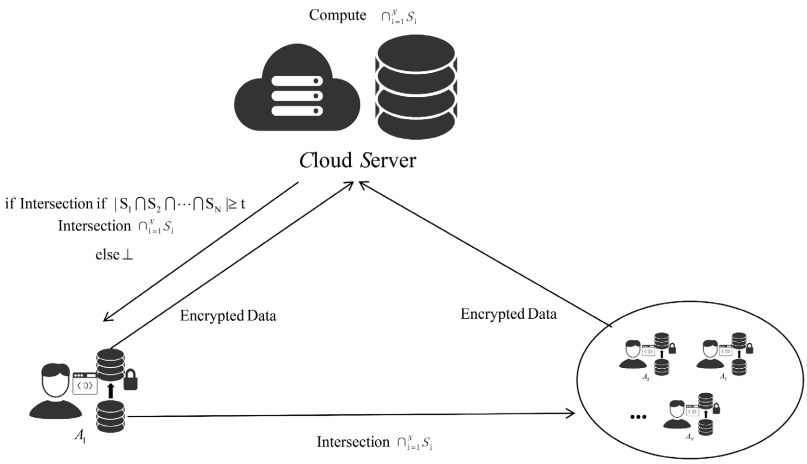

Figure 1. Traditional MPSI system model
图 1. 传统 MPSI 系统模型

半诚实安全的 TG-MPSI

本文在半诚实模型下分析 TG-MPSI 的安全性。攻击者可腐化云服务器 C 或腐化部分参与方，但腐化实体在执行过程中均按协议发送消息，只是试图从交互视图中获取额外信息。为满足电商协作中“只泄露必要业务信号”的要求，本文主要关注如下安全目标：

输入隐私：除协议允许泄露的输出(触发标志 $flag$ 与在 $flag=1$ 时的交集 R)外，任意被腐化实体不应获得关于诚实方输入集合 S_i 的额外信息(例如非交集元素、集合规模细节、结构性特征等)。

输出正确性：协议输出应与理想功能一致，即当满足触发条件时输出真实交集，否则输出 $\perp$ 。

阈值判定最小披露：阶段 I 仅允许输出触发判定所必需的信息(如 $flag$ 或与阈值比较等价的最小统计量)，避免在稀疏交集下暴露不必要的中间结构与路径信息。

TG-MPSI 系统模型如图 2 所示。

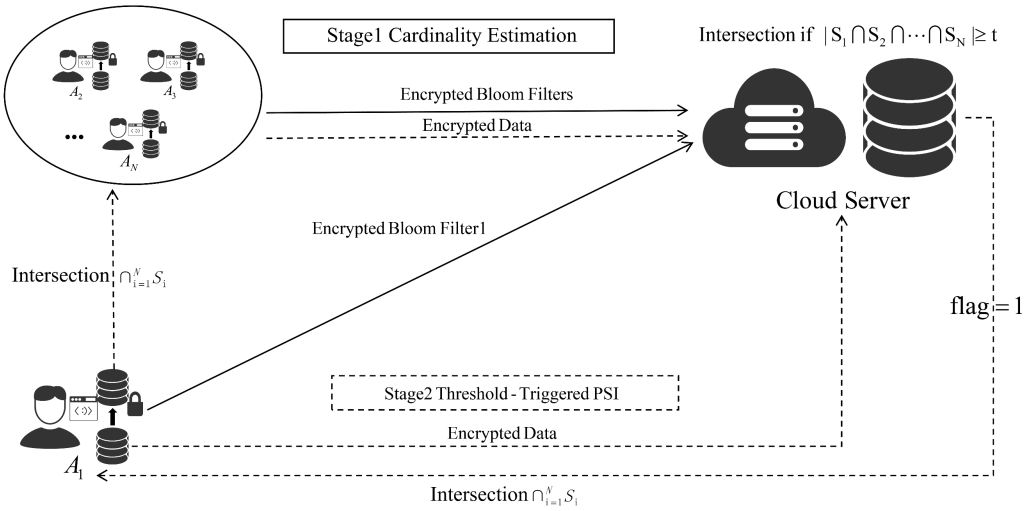

Figure 2. TG-MPSI system model diagram
图 2. TG-MPSI 系统模型图

给定协议 π_{TG} 与理想功能 $F_{TG0MPSI}$，若对任意 PPT 半诚实攻击者 A，存在 PPT 模拟器 Sim，使得对任意输入向量 S 有：

$$Real_{\pi_{TG},A}(\lambda,S) \approx_c Ideal_{F_{TG-MPSI},Sim,A}(\lambda,S) \tag{4}$$

则称协议 π_{TG} 在半诚实模型下安全。其中 $\approx$ 表示计算不可区分，即对任意 PPT 区分器 D：

$$\left|Pr\left[A(Real)=1\right]-Pr\left[A(Ideal)=1\right]\right| \leq negl(\lambda) \tag{5}$$

该定义保证，在电商协作场景中，除“是否触发深度协作”与“触发后的交集结果”之外，云端或被腐化参与方无法从协议交互中推断更多关于其他平台风险标识集合的内容。

5. 阈值触发多方隐私集合求交

本节介绍 TG-MPSI 在半诚实模型下的协议流程。方案面向电子商务跨域协作：多平台在不共享原始名单的前提下，对风控标识集合进行交叉核验。考虑到联防并非每次都需输出精确交集，TG-MPSI 采用两阶段门控设计：阶段 I 在云端进行隐私保护的交集规模估计，仅输出触发标志 $flag$ (可选输出估计值)以判断是否启动深度协作；阶段 II 仅在 $flag=1$ (估计交集规模达到业务阈值 T)时执行 MP-PSI 输出真实

交集，用于联合封禁、联合拦截或营销去重触达等联动处置。

5.1. 阶段 I：交集基数门控

阶段 I 的目标是判断真实交集规模 $\left|\bigcap_{i=1}^{n} S_i\right|$ 是否达到阈值 T，并输出触发标志 $flag \in 0,1$，同时尽量不泄露除触发判定所需之外的任何信息。

(1) Bloom 结构构造与加密上传

各参与方 P_i 选取 k 个公开哈希函数 $H = h_{\ell}{}_{\ell=1}^{k}$，对其本地风险标识集合 S_i 中的元素 x_j^i 计算哈希位置：

$$\left\{h_1\left(x_j^i\right), h_2\left(x_j^i\right), \cdots, h_k\left(x_j^i\right)\right\} \subseteq \{1, \cdots, m\} \tag{6}$$

初始化比特向量 $BF_i \in 0,1^m$，并对每个 $x \in S_i$ 置位 $BF_i\left[h_{\ell}\left(x\right)\right] \leftarrow 1$。随后，Leader 生成 Paillier 密钥对：

$$\left(pk, sk\right) \leftarrow \text{Paillier.KeyGen}\left(1^{\lambda}\right) \tag{7}$$

并将公钥 pk 分发给所有参与方与云服务器 C，私钥 sk 仅由 Leader 持有。参与方对每个槽位加密：

$$C_{i,j} = Enc_{pk} BF_i\left[j\right] \quad j = 1, \cdots, m \tag{8}$$

并将密文向量 $C_i = \left(C_{i,1}, \cdots, C_{i,m}\right)$ 上传至云服务器。

(2) 云端同态聚合与“全 1 槽位”构造

云服务器对每个槽位 j 在密文域做同态加和聚合：

$$C_j = \prod_{i=1}^{N} C_{i,j} = Enc_{pk}\left(\sum_{i=1}^{N} BF_i\left[j\right]\right) \tag{9}$$

为判定槽位 j 是否为全 1 插槽，云端进一步构造：

$$D_j = C_j \cdot Enc_{pk}\left(-n\right) = Enc_{pk}\left(\sum_{i=1}^{N} BF_i\left[j\right] - n\right) \tag{10}$$

将 $D_{j}{}_{j=1}^{m}$ 发送给 Leader 解密。

(3) Leader 解密判定与基数估计

Leader 解密得到每个槽位的明文差值并判定是否“全 1”：

$$F_j = 1\left[Dec\left(D_i\right) = 0\right] \tag{11}$$

其中 $F_j = 1$ 表示第 j 个槽位在所有参与方中均为 1。令全 1 槽位总数为

$$Z = \sum_{j=1}^{m} F_j \tag{12}$$

基于 Bloom 过滤器的标准估计式，可得到交集基数估计值：

$$\hat{n}_{\cap} = -\frac{m}{k} \ln\left(1 - \frac{Z}{m}\right) \tag{13}$$

最后进行阈值门控，若 $\hat{n}_{\cap} \geq T$，则置 $flag = 1$ 并进入阶段 II；否则置 $flag = 0$，协议终止并输出 $\perp$。

5.2. 阶段 II：触发后精确求交

当且仅当 $flag = 1$ 时，执行阶段 II 输出真实交集元素，用于跨平台协同风控处置或营销去重。

(1) 离线阶段：GBF 构造与混淆

各参与方基于其集合 S_i 构造混淆布隆过滤器 $GBF_i\left[1..m\right]$。对每个元素 $x \in S_i$，计算索引集合

$I_x = h_1(x), \cdots, h_k(x)$，并生成 k 份 XOR 份额写入对应槽位，使得满足：

$$\bigoplus_{j \in I_x} GBF_i[j] = x \tag{14}$$

对未赋值的槽位用随机串填充。为隐藏结构与访问模式，可进一步采用随机化/路径隐藏等混淆策略，使云端难以从槽位分布推断元素位置。

(2) 在线阶段：云端 XOR 聚合与 OT 受控解码

云服务器接收所有 GBF_i 后，对每个槽位做按位 XOR 聚合：

$$GBF_{\cap}[j] = GBF_1[j] \oplus GBF_2[j] \oplus \cdots \oplus GBF_m[j] \tag{15}$$

随后，Leader 与云端执行 1-out-of-2 OT 机制，使 Leader 只能按其本地索引集合查询必要槽位而不暴露查询位置。对任意候选元素 $x \in S_\ell$ (Leader 本地集合)，若满足：

$$GBF_{\cap}\left[h_1(x_j)\right] \oplus GBF_{\cap}\left[h_2(x_j)\right] \oplus \cdots \oplus GBF_{\cap}\left[h_k(x_j)\right] = 0^\lambda \tag{16}$$

则判定 $x \in \bigcap_{i=1}^{n} S_i$，并输出为交集元素。Leader 汇总交集结果并分发给其余参与方，用于电商联防/协同治理任务。

5.3. 正确性分析

定理 1 (阈值判定的正确性)

在 Bloom 参数 (m,k) 合理配置、误差概率受控的条件下，阶段 I 基于全 1 槽位数 Z 计算得到的 $\hat{n}_{\cap}$ 能以高概率正确反映交集规模是否超过阈值 T。当真实交集规模较小(典型电商稀疏重叠场景)时，方案以高概率输出 $flag = 0$ 并提前终止；当交集规模达到业务阈值时，以高概率输出 $flag = 1$ 触发阶段 II。

记真实交集规模为 $n_{\cap}$，则全 1 槽位计数 $Z = \sum_{j=1}^{m} F_j$ 的期望近似为

$$\mu = E[Z] = m\left(1 - \mathrm{e}^{-kn_{\cap}/m}\right) \tag{17}$$

阈值 T 在 Z 域对应：

$$Z_T = m\left(1 - \mathrm{e}^{-kT/m}\right) \tag{18}$$

在独立哈希近似下，Z 满足 Chernoff 集中界，从而第一类错误(误触发)与第二类错误(漏触发)分别满足

$$Pr[FP] = Pr\left[Z \geq Z_T \mid n_{\cap} < T\right] \leq \exp\left(-\frac{\mu}{3}\left(\frac{Z_T}{\mu} - 1\right)^2\right) \tag{19}$$

$$Pr[FN] = Pr\left[Z \geq Z_T \mid n_{\cap} < T\right] \leq \exp\left(-\frac{\mu}{2}\left(1 - \frac{Z_T}{\mu}\right)^2\right) \tag{20}$$

式(19)~(20)表明：当 $n_{\cap}$ 远离 T 时误判概率随 μ 指数下降；当 $n_{\cap} \approx T$ 时误判更敏感，需通过增大 m、合理选择 k 或设置阈值提升门控稳定性。

定理 2 (阶段 II 交集恢复的正确性)

当 $flag = 1$ 时，所有参与方对交集元素的 GBF 份额在云端 XOR 聚合后满足式(16)的零判定条件，从而 Leader 可正确恢复真实交集 $\bigcap_{i=1}^{n} S_i$。当 $flag = 0$ 时，协议不进入阶段 II，输出 $\perp$，符合门控语义。

5.4. 安全性分析

本节在第 3 章安全模型基础上，说明 TG-MPSI 在静态半诚实对手下满足输入隐私与最小披露原则，适用于电商多方协作的合规与商业敏感约束。

(1) 阶段 I：密文聚合的隐私性

仅 Leader 持有 Paillier 私钥 sk，云服务器与其他参与方无法解密任意 $C_{i,j}$、C_j、D_j。云端观察到的仅为 Paillier 密文，结合 IND-CPA 安全性，无法推断任意参与方 Bloom 位向量的具体分布，更无法恢复电商风险标识集合内容。

(2) 最小披露与门控输出

阶段 I 对外仅公开触发标志 $flag$ (以及可选的估计值 $\hat{n}_{\cap}$)。当 $flag=0$ 时协议终止，避免在“交集很小但仍强行做 PSI”时产生的额外交互与潜在侧信道泄露；这与电商风控链路“低价值不联动”的业务逻辑一致。

(3) 阶段 II：GBF 混淆 + OT 的访问模式保护

在阶段 II 中，云端仅看到经混淆的 GBF 槽位与 XOR 聚合结果；Leader 的槽位查询通过 1-out-of-2 OT 完成，使云端难以获知 Leader 具体查询的槽位位置，从而隐藏访问模式，降低由查询路径推断元素的风险。

定理 3 (半诚实安全性)

若 Paillier 满足 IND-CPA 安全性，OT 协议在半诚实模型下安全，哈希函数与 PRF 满足伪随机性，且 GBF 混淆使单方视图与随机串不可区分，则 TG-MPSI 在第 3 章定义的半诚实模型下安全，为进一步降低 Z 与 $\hat{n}_{\cap}$ 被单点掌握带来的合谋风险，可将阶段 I 的 Paillier 解密从“Leader 单点解密”升级为 (t,n) 阈值解密，即使 Leader 与云服务器 C 合谋，只要其腐化的解密份额数小于 t，除允许泄露的 $flag$ 与在 $flag=1$ 时的交集结果外，攻击者无法获得关于诚实方输入集合及 $(Z,\hat{n}_{\cap})$ 的额外信息。

6. 工程化落地

将参与方平台记为 $P_1,\cdots,P_n$，云侧聚合服务记为 C。各平台仅上传经加密、混淆后的 Bloom/GBF 结构，禁止上传原始风险名单；密钥与解密能力集中于 Leader (或合规托管模块)，云端仅执行密文聚合与结构化计算。对异常请求(高频触发、重复提交、参数越界)进行限流与风控拦截，避免通过调用频率推断业务策略。

对关键操作，如：KeyGen、上传、聚合、解密、触发判定、OT 交互，记录不可抵赖审计日志，包含请求 ID、时间戳、参数摘要与签名链路；对 flag 触发与交集下发设置双人复核或策略审批，支持事后追责与合规核查。必要时将日志摘要周期性锚定到独立存储以提升防篡改能力。

默认仅输出触发标志 $flag$；仅在 $flag=1$ 且满足业务授权策略时输出交集元素。交集结果以“最小字段集”形式下发(如标准化 ID/哈希标识)，并采用分级权限控制与到期回收机制；对跨境/跨主体协作场景，按数据最小化与目的限定原则配置阈值 T 与输出粒度，确保“可用、可控、可审计”。

7. 结论及展望

本文面向电商跨平台风控与协同治理中“交集稀疏但频繁比对”的需求，针对传统多方 PSI 无条件执行带来的通信与在线时延冗余及隐私暴露问题，提出阈值门控 TG-MPSI。方案在云端对加密 Bloom 过滤器进行安全聚合以估计交集基数，仅输出 1 比特触发标志；仅当估计规模达到阈值时，才启动基于 GBF 与 OT 的多方 PSI 输出交集用于联动处置，实现按需协作。在半诚实模型下，方案满足输入隐私与输出正确性。实验表明，在不同参与方数量、集合规模与阈值配置下，TG-MPSI 可在大量未触发场景显著降

低总通信开销与在线延迟，适用于电商风控联防、营销去重与商家治理等业务链路。未来可扩展至恶意安全与可验证计算，引入更细粒度门控与功能负载，并结合更轻量的基数估计、TEE 或差分隐私提升鲁棒性与可用性。

基金项目

国家自然科学基金联合基金重点项目(U24A20241)；贵州省重大专项(黔科合重大专项字[2024]014，黔科合重大专项字[2024]003)；现代商贸深度融合新零售电商数字经济平台关键技术研究(合同编号：黔科合支撑[2023]一般 231)。

参考文献

[1] Meadows, C. (1986) A More Efficient Cryptographic Matchmaking Protocol for Use in the Absence of a Continuously Available Third Party. 1986 *IEEE Symposium on Security and Privacy*, Oakland, CA, 7-9 April 1986, 134-137. https://doi.org/10.1109/sp.1986.10022

[2] 中国信息通信研究院安全研究所, 阿里巴巴集团安全部, 北京数牍科技有限公司. 隐私保护计算技术研究报告(2020 年) [R]. 北京: 中国信息通信研究院, 2020.

[3] 李凤华, 李晖, 牛犇, 陈金俊. 隐私计算——概念、计算框架及其未来发展趋势[J]. 工程(英文), 2019, 5(6): 1179-1192.

[4] 高莹, 王玮. 多方隐私集合交集计算技术综述[J]. 电子与信息学报, 2023, 45(5): 1859-1872.

[5] Feng, L., Liu, Y., Hu, K., Zeng, X., Fang, F., Xie, J., *et al*. (2024) LPP-BPSI: A Location Privacy-Preserving Scheme Using Blockchain and Private Set Intersection in Spatial Crowdsourcing. *Future Generation Computer Systems*, **157**, 112-123. https://doi.org/10.1016/j.future.2024.03.036

[6] Dörre, F., Mechler, J. and Müller-Quade, J. (2023) Practically Efficient Private Set Intersection from Trusted Hardware with Side-Channels. In: *Lecture Notes in Computer Science*, Springer, 268-301. https://doi.org/10.1007/978-981-99-8730-6_9

[7] Abadi, A., Dong, C., Murdoch, S.J. and Terzis, S. (2022) Multi-Party Updatable Delegated Private Set Intersection. In: *Lecture Notes in Computer Science*, Springer, 100-119. https://doi.org/10.1007/978-3-031-18283-9_6

[8] 张静, 罗守山, 杨义先, 辛阳. 安全两方集合交集云外包计算协议[J]. 北京邮电大学学报, 2019, 42(2): 13-18.

[9] 魏立斐, 刘纪海, 张蕾, 王勤, 贺崇德. 面向隐私保护的集合交集计算综述[J]. 计算机研究与发展, 2022, 59(8): 1782-1799.

[10] 李晴雯, 关海梅, 李晖. 面向应用的隐私集合求交技术综述[J]. 网络空间安全科学学报, 2023, 1(3): 68-85.

[11] Pinkas, B., Schneider, T. and Zohner, M. (2018) Scalable Private Set Intersection Based on OT Extension. *ACM Transactions on Privacy and Security*, **21**, 1-35. https://doi.org/10.1145/3154794

[12] Yang, Y., Dong, X., Cao, Z., Shen, J., Li, R., Yang, Y., *et al*. (2024) EMPSI: Efficient Multiparty Private Set Intersection (with Cardinality). *Frontiers of Computer Science*, **18**, Article No. 181804. https://doi.org/10.1007/s11704-022-2269-0

[13] 贾轩, 白玉真, 马智华. 隐私计算应用场景综述[J]. 信息通信技术与政策, 2022, 48(5): 45-52.

[14] 吴向阳. 大数据技术对跨境电商的影响[J]. 商场现代化, 2025(20): 46-48.

[15] 陈业旭. 电商平台隐私保护和销售策略问题研究综述[J]. 中国储运, 2025(5): 161.

[16] Kissner, L. and Song, D. (2005) Private and Threshold Set-Intersection. Technical Report CMU-CS-05-113. Carnegie Mellon University.

[17] Hu, J., Zhao, Y., Tan, B.H.M., Aung, K.M.M. and Wang, H. (2024) Enabling Threshold Functionality for Private Set Intersection Protocols in Cloud Computing. *IEEE Transactions on Information Forensics and Security*, **19**, 6184-6196. https://doi.org/10.1109/tifs.2024.3402355

[18] Dong, C., Chen, L. and Wen, Z. (2013) When Private Set Intersection Meets Big Data: An Efficient and Scalable Protocol. *Proceedings of the* 2013 *ACM SIGSAC Conference on Computer & Communications Security*, Berlin, 4-8 November 2013, 789-800.