基于双曲正切迭代的高效同态密文比较方案

杨 涛[1] 李星宇[2] 王海程[2] 纵 绚[3*]
YANG Tao LI Xingyu WANG Haicheng ZONG Xuan

摘 要 随着云计算的普及，在不可信环境中对海量敏感数据进行隐私保护处理的需求日益迫切。现有方法在追求高精度时往往导致计算深度过大和效率低下。针对此问题，文章提出了一种基于双曲正切函数（tanh）迭代的高效同态比较方案。该方法的核心思路是采用“分而治之”的策略。首先，通过大幅缩放输入，在极小值域内使用一个极低次多项式对 $\tanh(kx/2^N)$ 进行高精度初始近似。后巧妙利用 tanh 的倍角公式 $\tanh(2z)=2\tanh(z)/(1+\tanh^2(z))$ 进行 N 次迭代恢复。其中，每次迭代中分母的逆运算 $(1+\tanh^2(z))^{-1}$ 作用于一个固定的、极小的区间 [1,2)，因而可以被一个低次多项式高效逼近。通过将复杂的全局逼近分解为一系列基于低次多项式的、结构化的局部运算，所提方案有望在保证高逼近精度的同时，显著降低同态比较的乘法深度与总计算量。理论分析与模拟实验结果表明，该方法相较于现有技术在效率和精度方面具有明显优势，为推动同态搜索及更广泛的隐私保护数据分析应用的实用化提供了有效途径。

关键词 双曲正切；同态加密；多项式逼近；双曲正切函数

doi：10.3969/j.issn.1672-9528.2025.09.035

0 引言

随着数字化浪潮席卷全球，数据的集中存储与处理成为常态[1-2]。企业与个人用户日益依赖云服务提供商进行数据管理，以降低成本、提升效率[3]。然而，将敏感数据存储于第三方云平台，引发了对数据隐私和安全的深切忧虑[4]。传统的加密手段虽然能在数据传输和静态存储时提供保护，但在数据需要被云端分析或检索时，往往要求先解密，这无疑将数据暴露于潜在的不可信环境中。全同态加密[5]（fully homomorphic encryption, FHE）作为一种前沿的密码学范式，允许在加密状态下对数据执行任意复杂的计算，其计算结果解密后与在明文上执行相同运算的结果一致[6-7]。这一特性为解决上述矛盾提供了根本性的理论途径，对构建可信数据处理与共享生态具有里程碑式的意义[8]。

本文聚焦于对双曲正切函数 $\tanh(x)$ 这一经典的 S 型函数进行深度挖掘和优化利用[9]。然而，直接对 $\tanh(kx)$ 进行全局高精度多项式逼近依然面临困难。本文的核心思路是采用一种“分而治之”的迭代策略：首先，通过将输入 kx 大幅缩放至一个极小的值域 $kx/2^N$，在这个极小值域内，$\tanh(kx/2^N)$ 可以被一个次数非常低的多项式 P_{init} 以极高的精度进行初始近似。随后，利用双曲正切函数的倍角公式进行 N 次迭代。在每次迭代中，分母的逆运算由于其输入 $\tanh^2(z)$ 的范围固定在 [0,1)，也可以被一个针对此特定小区间的低次多项式 P_{inv} 高效近似。通过这种结构化的迭代，期望将复杂的全局逼近问题分解为一系列更易于处理的、基于低次多项式的局部运算，从而在保证最终逼近精度的前提下，有效控制同态计算的乘法深度和总计算量。

1 基于双曲正切函数的高效同态符号函数近似方法

1.1 $\tanh(kx)$ 近似符号函数

双曲正切函数 $\tanh(z)$ 定义为：

$$\tanh(z)=\frac{e^z-e^{\ z}}{e^z+e^{-z}}=\frac{e^{2z}-1}{e^{2z}+1} \tag{1}$$

$\tanh(kx)$ 作为 $\text{sign}(x)$ 的一个良好光滑近似，目标是将 $\tanh(kx)$ 表示为在同态加密环境下高效计算的多项式形式[10]。

1.2 泰勒展开 $\tanh(z)$ 为多项式形式

为了在仅支持加法和乘法的同态加密[11-12]环境中计算 $\tanh(z)$，直接方法是利用其在 $z=0$ 处的麦克劳林级数：

$$\tanh(z)=\sum_{n=1}^{\infty}\frac{B_{2n}4^n(4^n-1)}{(2n)!}z^{2n-1} \tag{2}$$

$$P_m(z)=c_1z+c_3z^3+\ldots+c_{2m-1}z^{2m-1} \tag{3}$$

1. 滨州市科技创新发展研究院 山东滨州 256606
2. 滨州职业学院 山东滨州 256603
3. 浙江交通职业技术学院 浙江杭州 311112
[基金项目] 中国高校产学研创新基金（2023IT104）

1.3 设计利用倍角公式进行优化

为了克服泰勒展开在较大输入值时精度不足的问题，同时控制多项式的计算复杂度，采用基于双曲正切函数的倍角公式进行迭代计算[13]。双曲正切函数的倍角公式为：

$$\tanh(2z)=\frac{2\tanh(z)}{1+\tanh^2(z)} \tag{4}$$

该公式从 tanh(z) 的值计算出 tanh(2z) 的值。算法 1 为基于倍角公式优化计算 tanh(kx) 近似值的流程。

算法 1：基于倍角公式的 tanh(kx) 近似算法

输入：原始值 x；缩放因子 k；迭代次数 N；

初始近似多项式 $P_{\text{init}}(z)$；逆元近似多项式 $P_{\text{inv}}(Y)$。

输出：tanh(kx) 的近似值 t_N。

1. 计算缩放后的初始输入：$s_0 \leftarrow kx/2^N$

2. 计算初始近似值：$t_0 \leftarrow P_{\text{init}}(s_0)$

3. for i=1 to N do

4. 计算分母项：$Y_{i-1} \leftarrow 1+t_{i-1}^2$

5. 计算分母项的逆元近似：$\text{inv\_}Y_{i-1} \leftarrow P_{\text{inv}}(Y_{i-1})$

6. 更新近似值：$t_i \leftarrow 2 \cdot t_{i-1} \cdot \text{inv\_}Y_{i-1}$

7. end for

8. return t_N

1.4 同态加密实现（基于 CKKS 方案）

上述算法 ApproxTanhIterative 由多项式运算构成，因此可以直接转化为在同态加密方案上的密文运算[14]。

（1）参数选择

k 和 N 的选择需要权衡逼近精度和计算开销。$P_{\text{init}}(z)$ 和 $P_{\text{inv}}(Y)$ 的次数和系数需要精心选择，以在其各自输入区间内达到所需的逼近精度，同时控制其在 CKKS 中计算的乘法深度[15-16]。

（2）CKKS 方案下的运算

算法中的所有变量 (x, s_0, t_i, Y_i, inv_Y_i) 均以 CKKS 密文形式存在。

加法对应 CKKS 的同态加法“EvalAdd”。

乘法对应 CKKS 的同态乘法“EvalMult”。

CKKS 的同态常数乘法“EvalMultByConst”或“EvalMultPlain”。

多项式求值 $P_{\text{init}}(s_0)$ 和 $P_{\text{inv}}(Y_{i-1})$ 可以使用如 Paterson-Stockmeyer 算法或简单的霍纳法则的同态版本来实现。

2 理论分析

2.1 精度和误差分析

2.1.1 初始近似误差ϵ_{init}

令 $s_0=kx/2^N$，初始近似误差定义为：

$$\epsilon_{\text{init}}(s_0)=P_{\text{init}}(s_0)-\tanh(s_0) \tag{5}$$

2.1.2 逆元近似误差ϵ_{inv}

在每次迭代中，使用 $P_{\text{inv}}(Y_{i-1})$ 来近似 $1/Y_{i-1}$，其中 $Y_{i-1}=1+t_{i-1}^2$。令 $t_{i-1}^*=\tanh(S/2^{N-i+1})$ 为理想的中间值。理想的分母项为 $Y^*_{i-1}=1+(t^*_{i-1})^2$。由于 t_{i-1} 本身是 t^*_{i-1} 的近似，所以 Y_{i-1} 也是 Y^*_{i-1} 的近似。

2.1.3 迭代误差传播

令 t_i^* 表示第 i 次迭代后 $\tanh(S/2^{N-i})$ 的真实值，而 t_i 是算法计算得到的近似值。定义第 i 步的累积误差为 $\Delta_i=t_i-t_i^*$。$t_i=2t_{i-1}\cdot P_{\text{inv}}(1+t_{i-1}^2)$。因此，用公式表示为：

$$\Delta_i=2t_{i-1}P_{\text{inv}}(1+t_{i-1}^2)-\frac{2t_{i-1}^*}{1+(t_{i-1}^*)^2} \tag{6}$$

进一步，$t_{i-1}=t_{i-1}^*+\Delta_{i-1}$。

代入并进行一阶近似分析：

$$t_i \approx t_i^*+\Delta_{i-1}\cdot\frac{2\left(1-(t_{i-1}^*)^2\right)}{\left(1+(t_i^*{}_1)^2\right)^2}+2t_{i-1}^*\epsilon_{\text{inv},i-1} \tag{7}$$

Δ_i 大致可以表示为前一步误差 Δ_{i-1} 的线性放大，加上由逆元近似引入的误差项。

$$|\Delta_i| \lesssim J(t_{i-1}^*)|\Delta_{i-1}|+2|t_{i-1}^*|\delta_{\text{inv}} \tag{8}$$

如果 J_i 是第 i 步的误差放大因子，总误差的界可以大致表示为：

$$|\Delta_N| \approx \left(\prod_{j=0}^{N-1}J_j\right)\delta_{\text{init}}+\sum_{l=0}^{N-1}\left(\prod_{j=l+1}^{N-1}J_j\right)\cdot(2|t_l^*|\delta_{\text{inv}}) \tag{9}$$

2.2 计算复杂度

2.2.1 乘法次数

（1）计算 $s_0=kx/2^N$。

（2）评估 $P_{\text{init}}(s_0)$。

（3）N 次迭代，每次迭代计算 t_{i-1}^2，评估 $P_{\text{inv}}(Y_{i-1})$ 计算 $2\cdot t_{i-1}\cdot\text{inv\_}Y_i$。

因此，总的密文 - 密文乘法次数 M_{total} 约为：

$$M_{\text{total}} \approx M_{\text{init}}+N\cdot(1+M_{\text{inv}}+1)=M_{\text{init}}+N(M_{\text{inv}}+2) \tag{10}$$

2.2.2 乘法深度

（1）评估 $P_{\text{init}}(s_0)$：若使用霍纳法则[17]，深度约为 $L_{\text{init}} \approx \lceil\log_2(d_{\text{init}})\rceil$ 或 $d_{\text{init}}/2$。

（2）N 次迭代，一次迭代的深度 $L_{\text{iter}} \approx 1(\text{for square})+L_{\text{inv}}+1(\text{for final mult})$。总的乘法深度约为：

$$L_{\text{total}} \approx L_{\text{init}}+N\cdot L_{\text{iter}}=L_{\text{init}}+N\cdot(L_{\text{inv}}+2) \tag{11}$$

3 实验结果与分析

3.1 实验设置

本章所有实验均基于 Python 环境下的模拟计算，并针对不同的目标精度 α 和死区参数 ϵ 进行调优，以期在合理的计

算开销下达到最佳逼近效果 [18]。

3.2 精度对比分析

3.2.1 不同输入值下的逼近误差

图 1 展示了在固定死区参数ϵ =0.05 的条件下，不同方法对符号函数的逼近误差随输入值 $x \in [-1, 1]$ 变化的曲线。

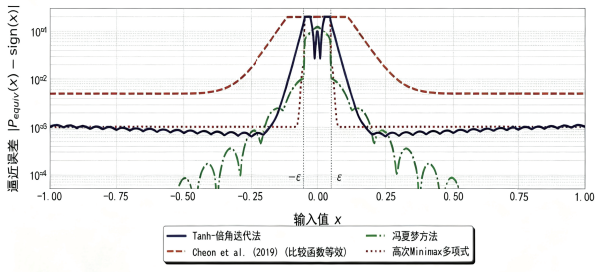

图 1 不同输入值下的逼近误差

从图 1 中可以观察到，本章提出的 tanh- 倍角迭代法在精度方面展现出显著优势：

（1）在有效工作区间 $|x| > \epsilon$ 内，tanh- 倍角迭代法的误差曲线不仅达到了所有比较方法中的最低水平，而且表现出高度的平坦性。

（2）相较于高次 Minimax 多项式在 ϵ 边界附近可能出现的吉布斯现象，以 Cheon (2019) 法 [19] 的比较函数等效方法在 x 接近零点时误差相对较大的情况，tanh- 倍角迭代法在死区边界处并未出现明显的误差峰值或剧烈波动。冯夏梦法 [20] 在误差的绝对大小和稳定性方面略逊于 tanh- 倍角迭代法。

3.2.2 最大误差对死区参数 ϵ 的敏感度

图 2 进一步探讨了当死区参数 ϵ 变化时，不同方法所能达到的最大逼近误差。ϵ 越小，对算法在接近零点处的逼近能力要求越高。anh- 倍角迭代法在应对小 ϵ 挑战时，随着 ϵ 从较大值向极小值变化，Tanh- 倍角迭代法的最大误差曲线增长最为平缓。即使在非常小的死区条件下，本方法依然能将最大误差控制在较低水平。

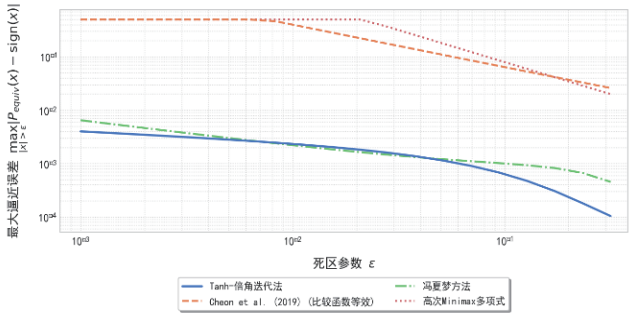

图 2 最大误差对死区参数 ϵ 的敏感度

3.3 效率对比分析

3.3.1 运行时间与目标精度的关系

图 3 对比了不同方法在达到不同目标精度比特数 α 时所需的模拟运行时间。随着目标精度 α 的增加，tanh- 倍角迭代法的运行时间增长最为平缓，表现出接近线性或低阶多项式的增长趋势。这与高次 Minimax 法的指数级时间增长形成了显著区别。当 α 值较大时，tanh- 倍角迭代法的模拟运行时间远低于其他方法。

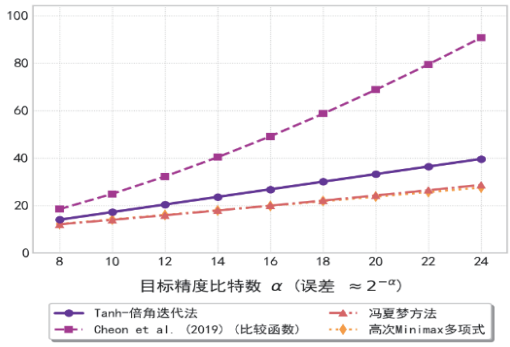

图 3 运行时间与目标精度的关系

3.3.2 乘法深度与目标精度的关系

乘法深度是同态加密应用中的一个硬性约束。图 4 比较了不同方法在达到目标精度 α 时所需的理论乘法深度。从图 4 中可以看出，tanh- 倍角迭代法在控制乘法深度方面，乘法深度曲线在所有比较方法中始终处于最低水平，并且随 α 的增加呈现平缓的增长。 即使在高精度要求下，tanh- 倍角迭代法也能将总计算深度维持在较低水平，这对减少同态加密方案的参数需求、控制噪声积累至关重要。

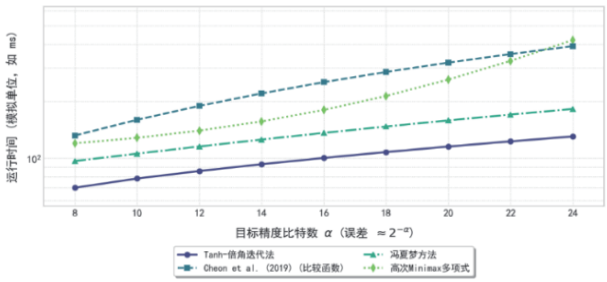

图 4 乘法深度与目标精度的关系

4 总结

本文针对同态加密环境下符号函数近似计算的效率与精度难题，提出了一种基于双曲正切函数（tanh）及其倍角公式迭代的优化方法。通过理论分析与模拟实验，验证了该方法的可行性与潜在优势。研究表明，本方法通过将初始逼近聚焦于输入值大幅缩放后的小范围，并结合固定区间的逆元多项式近似与结构化的倍角迭代，能够在保证高逼近精度的同时，有效控制计算开销，特别是在乘法深度和高精度需求下的运行时间增长方面，相较于部分现有方法展现出更优的理论特性。研究结果为在资源受限的同态加密应用中实现高效、精确的比较等关键逻辑运算提供了新的思路与技术途径，

对推动同态加密技术的实用化具有积极意义。

参考文献：

[1]MANYIKA J, CHUI M, BROWN B, et al. Big data: the next frontier for innovation,competition,and productivity[EB/OL].(2011-05-13)[2025-06-23].https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation.

[2]DEAN J, GHEMAWAT S. MapReduce: simplified data processing on large clusters[J]. Communications of the ACM, 2008, 51(1): 107-113.

[3]ARMBRUST M, FOX A, GRIFFITH R, et al. A view of cloud computing[J]. Communications of the ACM, 2010,53(4):50-58.

[4]SUBASHINI S, KAVITHA V. A survey on security issues in service delivery models of cloud computing[J]. Journal of network and computer applications, 2011, 34(1):1-11.

[5]GENTRY C. Fully homomorphic encryption using ideal lattices[C]//STOC'09: Proceedings of the forty-first annual ACM symposium on Theory of computing. NewYork:ACM,2009: 169-178.

[6]SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]// Proceeding 2000 IEEE Symposium on Security and Privacy. Piscataway:IEEE,2000:44-55.

[7]BONEH D, DI CRESCENZO G, OSTROVSKY R, et al. Public key encryption with keyword search[C]//Advances in Cryptology-EUROCRYPT 2004.Berlin:Springer,2004: 506-522.

[8]CHEON J H, KIM D, KIM D, et al. Numerical method for comparison on homomorphically encrypted numbers[C]// Advances in Cryptology-ASIACRYPT 2019: 25th International Conference on the Theory and Application of Cryptology and Information Security. NewYork: ACM, 2019: 415-445.

[9]NAEHRIG M, LAUTER K, VAIKUNTANATHAN V. Can homomorphic encryption be practical?[C]//Proceedings of the 3rd ACM workshop on Cloud computing security workshop. NewYork:ACM, 2011: 113-124.

[10]CURTMOLA R, GARAY J, KAMARA S, et al. Searchable symmetric encryption: improved definitions and efficient constructions[C]//Proceedings of the 13th ACM conference on Computer and communications security. NewYork:ACM,2006: 79-88.

[11]BETHENCOURT J, SAHAI A, WATERS B. Ciphertext-policy attribute-based encryption[C]//2007 IEEE symposium on security and privacy . Piscataway:IEEE, 2007: 321-334.

[12]KIM A, SONG Y, KIM M, et al. Logistic regression model training based on the approximate homomorphic encryption[J]. BMC medical genomics, 2018, 11: 23-31.

[13] 李媛 . 隐私保护下的不确定轨迹 k 近邻查询研究 [D]. 哈尔滨 : 哈尔滨理工大学 ,2022.

[14] 徐科鑫 , 王丽萍 . 多方全同态加密研究进展 [J]. 密码学报 (中英文),2024,11(4):719-739.

[15] 杨鸿健 , 胡学先 , 李可佳 , 等 . 隐私保护的非线性联邦支持向量机研究 [J]. 计算机科学 ,2022, 49(12):22-32.

[16]GENTRY C, HALEVI S, SMART N P. Homomorphic evaluation of the AES circuit[C]//Advances in Cryptology – CRYPTO 2012. Berlin: Springer, 2012: 850-867.

[17]LAUTER K, LÓPEZ-ALT A, NAEHRIG M. Private computation on encrypted genomic data[C]//Progress in Cryptology-LATINCRYPT 2014. Berlin: Springer, 2014: 3-27.

[18]GILAD-BACHRACH R, DOWLIN N, LAINE K, et al. CryptoNets: applying neural networks to encrypted data with high throughput and accuracy[C]//International conference on machine learning. NewYork:ACM, 2016: 201-210.

[19]CHEON J H, KIM D, KIM D, et al. Numerical method for comparison on homomorphically encrypted numbers[C]// International conference on the theory and application of cryptology and information security. Berlin: Springer, 2019: 415-445.

[20] 冯夏梦 . 基于同态加密的密文结构化数据搜索技术研究 [D]. 西安 : 西安电子科技大学 ,2024.

【作者简介】

杨涛（1972—），男，山东滨州人，硕士研究生，研究方向：计算机技术研究。

李星宇（1998—），女，山东淄博人，硕士研究生，研究方向：大数据、数据安全。

王海程（1999—），男，山东滨州人，硕士研究生，研究方向：大数据、数据安全。

纵绚（1991—），通信作者（email:zongxuan_zx@163.com），女，山东滨州人，博士研究生，研究方向：大数据、数据安全。

（收稿日期：2025-08-08 修回日期：2025-09-11）