与流行度识别),本方案所需时间开销明显低于 perfectDedup.此外,本方案摆脱了实时在线第三方 IS.因此,本方案在总时间开销上具有较明显的优势.

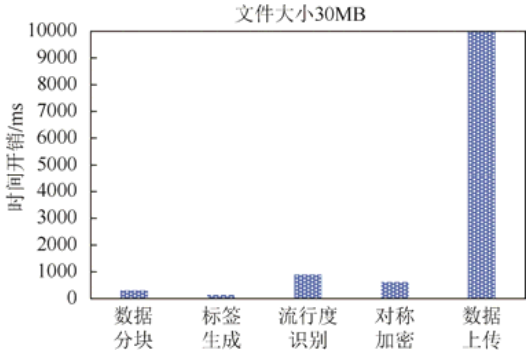

Fig.7 The time span for each phase of experiment ($Count_{FA}<T$)

图 7 实验中各阶段所需时间开销($Count_{FA}<T$)

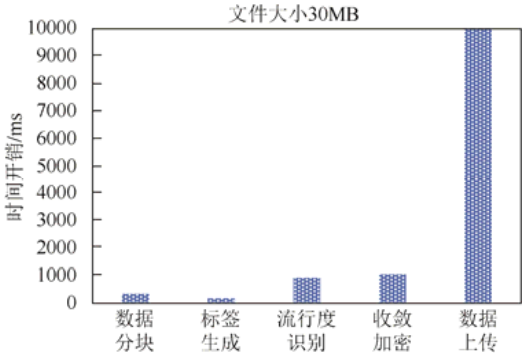

Fig.8 The time span for each phase of experiment ($Count_{FA}=T$)

图 8 实验中各阶段所需时间开销($Count_{FA}=T$)

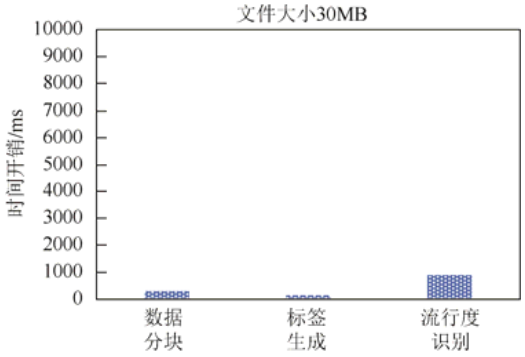

Fig.9 The time span for each phase of experiment ($Count_{FA}>T$)

图 9 实验中各阶段所需时间开销($Count_{FA}>T$)

Fig.10 Comparison of the time span between our scheme and the perfectDedup scheme ($Count_{FB}<T$)

图 10 本文方案与 perfectDedup 方案时间开销对比($Count_{FB}<T$)

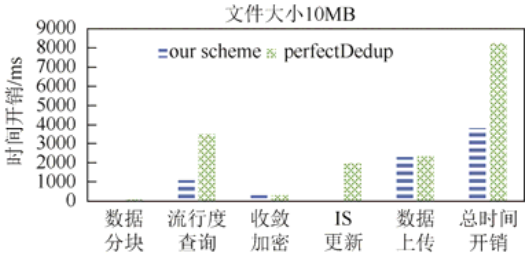

Fig.11 Comparison of the time span between our scheme and the perfectDedup scheme ($Count_{FB}=T$)

图 11 本文方案与 perfectDedup 方案时间开销对比($Count_{FB}=T$)

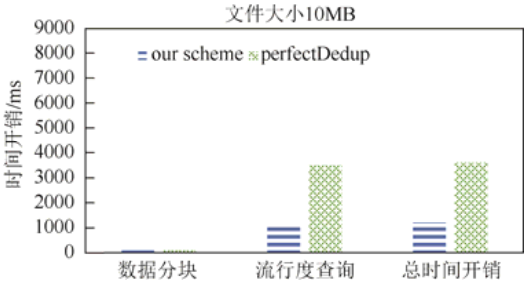

Fig.12 Comparison of the time span between our scheme and the perfectDedup scheme ($Count_{FB}>T$)

图 12 本文方案与 perfectDedup 方案时间开销对比($Count_{FB}>T$)

(3) 更少的存储空间开销

如图 13、图 14 所示,NoDedup 方案不执行重复数据删除,perfectDedup 方案无法删除非流行加密数据.本文方案的存储空间开销与持有数据的用户数量无关.且文件越大,本文方案的优势越明显.

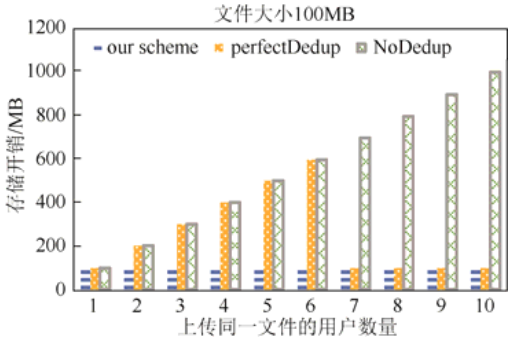

Fig.13 Cloud server storage costs of different schemes (Data size 100M)

图 13 3 种方案中云服务器存储开销对比(每个文件 100M)

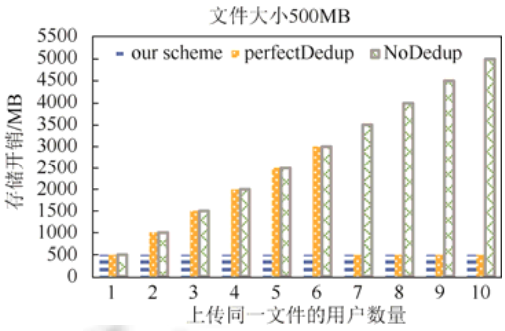

Fig.14 Cloud server storage costs of different schemes (Data size 500M)

图 14 3 种方案中云服务器存储开销对比(每个文件 500M)

(4) 性能分析比较

由以上实验结果可知,划分数据流行度、摆脱实时在线可信第三方能够显著提升重复数据删除方案的执行效率.表 1 给出了本文方案与其他代表性方案是否具有以上两个优点的分析和比较.

Table 1 Comparison of schemes characteristics

表 1 方案特点对比

方案	[7]	[8]	[9]	[15]	[17]	Our
划分数据流行度	×	√	√	×	×	√
摆脱实时在线可信第三方	×	×	×	×	√	√

7 总结与展望

本文研究了云存储环境下加密数据的重复删除问题,提出了一种基于离线密钥分发的加密数据重复删除方案.此方案通过构造语义安全的双线性映射,能够实现在不泄露数据任何明文信息的情况下完成流行度查询.通过广播加密为授权用户生成辅助密钥,保证非流行数据加密密钥的存储与传递的安全.持有相同非流行数据的不同用户能够获取相同的加密密钥,得到相同的加密数据,进而使云服务器能够对非流行数据进行重复数据删除.采用改进后的收敛加密算法保护隐私度较低的流行数据,用户能够自行生成加密密钥,进一步提高了方案的执行效率.通过安全分析与仿真实验,证明本方案具有较高的安全性与实用性.

如何摆脱广播中心,实现只有用户与云服务器两方交互的重复数据删除方案,是下一步的研究重点.

References:

[1] Fu YX, Luo SM, Shu JW. Survey of secure cloud storage system and key technologies. Journal of Computer Research and Development, 2013,50(1):136−145 (in Chinese with English abstract).

[2] Fu YJ, Xiao N, Liu F. Research and development on key techniques of data deduplication. Journal of Computer Research and Development, 2012,49(1):12−20 (in Chinese with English abstract).

[3] Ao L, Shu JW, Li MQ. Data deduplication techniques. Ruan Jian Xue Bao/Journal of Software, 2010,21(5):916−929 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3761.htm [doi: 10. 3724/SP.J.1001.2010.03761]

[4] Jeramiah B. Opendedup: Open-Source deduplication put to the test. Belltown Media, 2013. http://opendedup.org/

[5] Meyer DT, Bolosky WJ. A study of practical deduplication. ACM Trans. on Storage (TOS), 2012,7(4):14.

[6] Douceur JR, Adya A, Bolosky WJ, *et al*. Reclaiming space from duplicate files in aserverless distributed file system. In: Proc. of the ICDCS. IEEE, 2002. 617–624.

[7] Puzio P, Molva R, Onen M. Cloudedup: Secure deduplication with encrypted data for cloud storage. In: Proc. of the CloudCom. IEEE Computer Society, 2013. 363–370.

[8] Puzio P, Molva R, Önen M. PerfectDedup: Secure data deduplication. In: Proc. of the Int'l Workshop on Data Privacy Management. Springer Int'l Publishing, 2015. 150–166.

[9] Stanek J, Sorniotti A, Androulak E, *et al*. A secure data deduplication scheme for cloud storage. In: Christin N, Safavi-Naini R, eds. LNCS 8437. Springer-Verlag, 2014. 99–118.

[10] Xu J, Chang E C, Zhou J. Weak leakage-resilient client-side deduplication of encrypted data in cloud storage. In: Proc. of the ACM SIGSAC Symp. on Information, Computer and Communications Security. ACM, 2013. 195–206.

[11] Adya A, Bolosky WJ, Castro M, *et al*. Farsite: Federated, available, and reliable storage for an incompletely trusted environment. ACM SIGOPS Operating Systems Review, 2002,36(SI):1–14.

[12] Hur J, Koo D, Shin Y, *et al*. Secure data deduplication with dynamic ownership management in cloud storage. IEEE Trans. on Knowledge and Data Engineering, 2016,28(11):1.

[13] Perttula. Attacks on convergent encryption. 2008. https://tahoe-lafs.org/hacktahoelafs/drew_perttula.html

[14] Bellare M, Keelveedhi S, Ristenpart T. Message-Locked encryption and secure deduplication. In: Proc. of the EUROCRYPT. LNCS 7881, Springer-Verlag, 2013. 296–312.

[15] Mihir B, Keelveedhi S, Ristenpart T. DupLESS: Server-Aided encryption for deduplicated storage. In: Proc. of the 22nd USENIX Conf. on Security. USENIX Association, 2013. 179–194.

[16] Douceur JR. The Sybil attack. In: Proc. of the Peer-to-Peer Systems. Springer-Verlag, 2002. 251–260.

[17] Liu J, Asokan N, Pinkas B. Secure deduplication of encrypted data without additional servers. Technical Report, 455, ePrint archive, 2015. https://eprint.iacr.org/2015/455

[18] Li L, Xue R, Zhang HG, Feng DG, Wang L. Security analysis of authenticated key exchange protocol based on password. ACTA ELECTRONICA SINICA, 2005,33(1):166–170 (in Chinese with English abstract).

[19] Hu XX, Zhang ZF, Liu WF. Universal composable password authenticated key exchange protocol in the standard model. Ruan Jian Xue Bao/Journal of Software, 2011,22(11):2820–2832 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3910.htm [doi: 10.3724/SP.J.1001.2011.03910]

[20] Cui H, Deng RH, Li Y. Attribute-Based storage supporting secure deduplication of encrypted data in cloud. IEEE Trans. on Big Data, 2016, 1–13.

[21] Zhang XS. The construction and calculation of bilinear pairs in cryptography [Ph.D. Thesis]. Beijing: The Chinese Academy of Sciences, 2012 (in Chinese with English abstract).

[22] Chen YM, Cheng XG, Wang S. Pairing certificateless signature scheme based on information network security. Netinfo Security, 2017,(3):53–58 (in Chinese with English abstract).

[23] Sakai R, Furukawa J. Identity-Based broadcast encryption. Journal of Electronics & Information Technology, 2007,33(4): 1047–1050.

[24] Delerablée C. Identity-Based broadcast encryption with constant size ciphertexts and private keys. In: Proc. of the Advances in Crypotology, Int'l Conf. on Theory and Application of Cryptology and Information Security. Springer-Verlag, 2007. 200–215.

[25] Tan ZW, Liu ZJ, Xiao HG. A fully public key tracing and revocation scheme provably secure against adaptive adversary. Ruan Jian Xue Bao/Journal of Software, 2005,16(7):1333–1343 (in Chinese with English abstract). http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20050716&journal_id=jos [doi: 10.1360/jos161333]

[26] Pang LJ, Li HX, Jiao LC. Design and analysis of a provable secure multi-recipient public key encryption scheme. Ruan Jian Xue Bao/Journal of Software, 2009,20(10):2907–2914 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3552.htm [doi: 10.3724/SP.J.1001.2009.03552]

[27] Lynn B. The pairing-based cryptographic library. 2015. http://crypto.Stanford.edu/pbc/

[28] Loukides M, Oram A. Programming with GNU SoftWare. O'Reilly & Associates, 1997,86(3):350–359.

[29] Steiner M. The PBC_bce broadcast encryption library. 2006. https://crypto.stanford.edu/pbc/bce/

[30] Hu XT, Qin ZP, Zhang H, Hao GS. Research and improved implementation of AES algorithm in OpenSSL. Control & Automation, 2009,25(12):83–85.

附中文参考文献:

[1] 傅颖勋,罗圣美,舒继武.安全云存储系统与关键技术综述.计算机研究与发展,2013,50(1):136–145.

[2] 付印金,肖侬,刘芳.重复数据删除关键技术研究进展.计算机研究与发展,2012,49(1):12–20.

[3] 敖莉,舒继武,李明强.重复数据删除技术.软件学报,2010,21(5):916–929. http://www.jos.org.cn/1000-9825/3761.htm [doi: 10.3724/SP.J.1001.2010.03761]

[18] 李莉,薛锐,张焕国,冯登国,王丽娜.基于口令认证的密钥交换协议的安全性分析.电子学报,2005,33(1):166–170.

[19] 胡学先,张振峰,刘文芬.标准模型下通用可组合的口令认证密钥交换协议.软件学报,2011,22(11):2820–2832. http://www.jos.org.cn/1000-9825/3910.htm [doi: 10.3724/SP.J.1001.2011.03910]

[21] 张旭升,林东岱.密码学中双线性对的构造与计算[博士学位论文].北京:中国科学院大学,2012.

[22] 陈亚萌,程相国,王硕.基于双线性对的无证书群签名方案研究.信息网络安全,2017,(3):53–58.

[25] 谭作文,刘卓军,肖红光.一个安全公钥广播加密方案.软件学报,2005,16(7):1333–1343. http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20050716&journal_id=jos [doi: 10.1360/jos161333]

[26] 庞辽军,李慧贤,焦李成,王育民.可证明安全的多接收者公钥加密方案设计与分析.软件学报,2009,20(10):2907–2914. http://www.jos.org.cn/1000-9825/3552.htm [doi: 10.3724/SP.J.1001.2009.03552]

张曙光(1991－),男,山东曲阜人,硕士,主要研究领域为密码学,云计算安全.

刘红燕(1994－),女,硕士,主要研究领域为云中重复数据删除.

咸鹤群(1979－),男,博士,副教授,CCF 高级会员,主要研究领域为密码学,云计算安全,系统安全.

侯瑞涛(1993－),男,学士,主要研究领域为数据库数字水印.

王雅哲(1979－),男,博士,副研究员,主要研究领域为物联网安全,智能信息设备安全.