**参考文献**

【1】相似重复记录清理方法研究综述。叶焕倬 吴 迪

(中南财经政法大学信息与安全工程学院信息系 武汉 430073)

：56-57

【2】Galhardas H, F lorescu D, Sh asha D, et a.l Declarative Data

C lean ing: Language , M odel and A lgorithm s[ C ]. In: P roceedings

of the 27th International Conference on Very Large Data Bases,

R om e, Ita ly. 2001: 371- 380

【3】B itton D, D e W itt D J . Duplicate Record E li m in ation in Large Data

Files[ J]. ACM Transactions on Database System s, 1983, 8 ( 2 ):

255- 265.

【4】H ernandezM, Stolfo S. TheMerge /Pu rge Prob lem for LargeDatabases

[ C]. In: Proceedings of theACM SIGMOD InternationalConference on

Management of Data, S an Jose, California . 1995: 127- 138.

【5】H ernandezM A, Stolfo S J. Real- W orld Data is D irty : Data

C leansing and th e M erge /Purge Prob lem [ J]. D ata M ining and

K now ledge D iscovery, 1998, 2( 1): 9- 37.

【6】梁雪，任剑锋，景丽． 基于 ＱＰＳＯＬＳＳＶＭ 的数据库相似重复记录检

测算法［Ｊ］． 计算机科学，２０１２，３９（１ １ ）：１５７ １５９．

【7】马翔． 粒子群优化 ＢＰ 神经网络用于重复记录检测［ Ｊ］． 辽宁工程

技术大学学报：自然科学版，２０１０，２９（５）：９５９ ９６３．