Supplemental Material (Appendix)
SLocator: Localizing the Origin of SQL Queries in Database-Backed Web Applications

1 APPENDIX

This appendix contains the statistics of JPA API usage, workload, and SQL queries generated by the workload.

1.1 Statistics of JPA API usage in the studied applications

Table 1 shows the statistics of the JPA API usage in the studied applications. In total, there are 595 JPA calls in the studied applications, among which there are 293 basic CRUD JAP API calls (49.2%), 291 JPQL queries (48.9%), and 11 JPA criteria (1.8%). No native SQL queries are used in the studied applications.

TABLE 1: Statistics of the JPA API usage in the studied applications. No. of persist, find, merge, and remove refer to the number of basic CRUD JPA calls in the studied applications. No. of native SQL and JPQL refer to the number of native SQL queries and JPQL queries used in the studied applications, respectively.

Application	No. of persist	No. of find	No. of merge	No. of remove	No. of native SQL	No. of JPQL	No. of Criteria	Sum. of JPA APIs
PetClinic	3	1	3	0	0	5	0	12
CloudStore	11	5	0	1	0	23	0	40
WallRide	15	13	16	9	0	29	11	93
JeeWeb	39	13	13	26	0	21	0	112
PublicCMS	16	16	15	23	0	62	0	132
bbs	19	0	5	1	0	123	0	148
BroadleafCommerce	10	11	2	7	0	28	0	58
Sum. across applications	113	59	54	67	0	291	11	595

1.2 Coverage under the workload on the studied applications

The workload is generated by navigating the menu in application web pages and covers most user actions in each application (as discussed in Section 4.1). Table 2 shows the coverage of web requests, database accesses, and database tables under the workload on the studied applications. On average, the workload covers 71.2% of the related web requests, 70.7% of the related database accesses, and 74% of the related database tables. Note that, some web requests are not covered by the exercised workload because they are not executable via main user actions. For example, bbs provides the functionality to manage the third party login interfaces when a user cannot login successfully (i.e., an exception in the use case), which was not covered by the simulated workload.

TABLE 2: Coverage of web requests, database accesses, and database tables in the studied applications under the workload.

Application	Web requests	Database accesses	Database Tables
PetClinic	12/12 (100.0%)	12/12 (100.0%)	7/7 (100.0%)
CloudStore	12/12 (100.0%)	31/40 (77.5%)	10/11 (90.9%)
WallRide	36/56 (64.3%)	67/93 (72.0%)	17/32 (53.1%)
JeeWeb	31/45 (68.9%)	87/112 (77.7%)	21/31 (67.7%)
PublicCMS	32/62 (51.6%)	78/132 (59.1%)	22/43 (51.2%)
bbs	62/141 (44.9%)	77/148 (52.0%)	34/44 (77.3%)
BroadleafCommerce	33/48 (68.8%)	33/58 (56.9%)	28/36 (77.8%)
Avg. across applications	31/54 (71.2%)	55/85 (70.7%)	20/29 (74.0%)

1.3 Statistics of the distinct SQL queries

Table 3 shows the statistics of the distinct SQL queries generated by the workload (as discussed in Section 4.1). In total, there are 475 distinct SQL queries which cover 127 database tables. Among the SQL queries, there are 259 SELECT statements (54.5%), 79 INSERT statements (16.6%), 71 UPDATE statements (14.9%), and 66 DELETE statements (13.9%). 73 SQL queries (15.3%) use JOIN clauses while 7 SQL queries (1.5%) contain nested subselects. On average, each SQL query contains 1.7 tables.

TABLE 3: Statistics of the distinct SQL queries generated by the workload. SELECT, INSERT, UPDATE, and DELETE refer to corresponding SQL statements. joins and subselects refer to the SQL queries which use JOIN clauses and SQL queries which contain subselects, respectively.

Application	No. of queries	No. of tables	No. of SELECT	No. of INSERT	No. of UPDATE	No. of DELETE	No. of joins	No. of subselects	No. tables per query
PetClinic	13	7	8	3	2	0	5	0	1.6
CloudStore	31	10	20	7	3	1	11	0	1.7
WallRide	89	17	54	11	11	12	38	4	3.8
JeeWeb	123	21	80	15	12	16	7	3	1.2
PublicCMS	74	22	29	14	16	15	0	0	1.0
bbs	100	34	41	19	20	20	2	0	1.0
BroadleafCommerce	46	28	27	10	7	2	10	0	1.5
Sum. across applications	476	139	259	79	71	66	73	7	–
Pct. across applications	–	–	54.5%	16.6%	14.9%	13.9%	15.3%	1.5%	–

1.4 Statistics of the distinct slow SQL queries

Table 4 shows the statistics of the distinct slow SQL queries (as discussed in Section 4.3). In total, there are 48 distinct slow SQL queries which cover 56 database tables. Among the slow SQL queries, there are 34 SELECT statements (70.8%), 7 INSERT statements (14.6%), 1 UPDATE statements (2.1%), and 6 DELETE statements (12.3%). 16 slow SQL queries (33.3%) use JOIN clauses while 1 slow SQL queries (2.1%) contain nested subselects. On average, each slow SQL query contains 3.2 tables.

TABLE 4: Statistics of the distinct slow SQL queries generated by the workload. SELECT, INSERT, UPDATE, and DELETE refer to corresponding slow SQL statements. joins and subselects refer to the slow SQL queries which use JOIN clauses and slow SQL queries which contain subselects, respectively.

Application	No. of queries	No. of tables	No. of SELECT	No. of INSERT	No. of UPDATE	No. of DELETE	No. of joins	No. of subselects	No. tables per query
PetClinic	1	2	1	0	0	0	1	0	2.0
CloudStore	3	2	3	0	0	0	3	0	3.3
WallRide	9	15	9	0	0	0	9	1	12.3
JeeWeb	13	11	7	5	1	0	1	0	1.2
PublicCMS	7	7	3	0	0	4	0	0	1.0
bbs	10	13	7	1	0	2	1	0	1.1
BroadleafCommerce	5	6	4	1	0	0	1	0	1.2
Sum. across applications	48	56	34	7	1	6	16	1	–
Pct. across applications	–	–	70.8%	14.6%	2.1%	12.5%	33.3%	2.1%	–