



(51) International Patent Classification:

G06F 16/25 (2019.01) G06F 16/22 (2019.01)
G06F 16/2455 (2019.01) G06F 21/62 (2013.01)
G06F 16/28 (2019.01) H04L 65/40 (2022.01)

(21) International Application Number:

PCT/US2021/053840

(22) International Filing Date:

06 October 2021 (06.10.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/088,863 07 October 2020 (07.10.2020) US

(71) Applicant: **VISA INTERNATIONAL SERVICE ASSOCIATION** [US/US]; P.O. BOX 8999, SAN FRANCISCO, California 94128 (US).

(72) Inventors: **XU, Minghua**; PO BOX 8999, San Francisco, California 94128 (US). **CHRISTODORESCU, Mihai**; PO BOX 8999, San Francisco, California 94128 (US). **SUN, Wei**; PO BOX 8999, San Francisco, California 94128 (US). **RINDAL, Peter**; PO BOX 8999, San Francisco, California 94128 (US). **KUMARESAN, Ranjit**; PO BOX 8999, San Francisco, California 94128 (US). **NAGARAJA, Vinjith**; PO BOX 8999, San Francisco, California 94128 (US). **PATEL, Karankumar Hiteshbhai**; PO BOX 8999, San Francisco, California 94128 (US).

(74) Agent: **RACZKOWSKI, David B.** et al.; Kilpatrick, Townsend & Stockton LLP, 1100 Peachtree Street Suite 2800, Atlanta, Georgia 30309 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

(54) Title: SECURE AND SCALABLE PRIVATE SET INTERSECTION FOR LARGE DATASETS

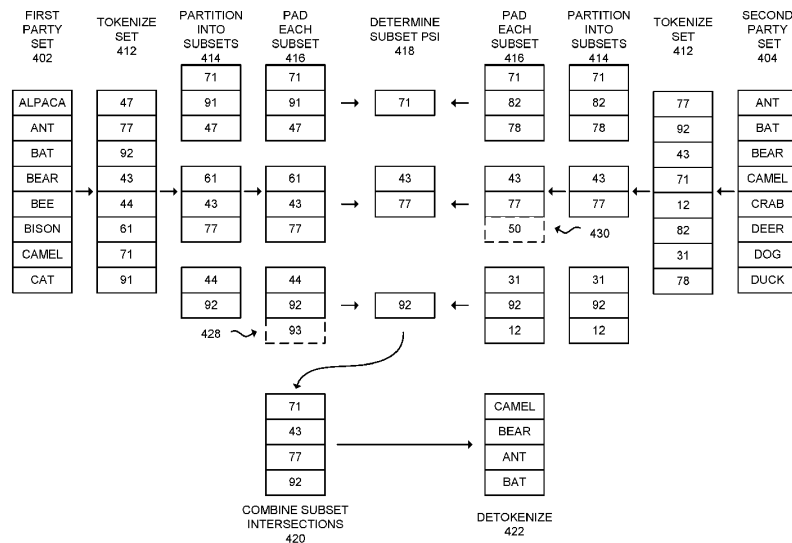


FIG. 4

(57) Abstract: Embodiments of the present disclosure are directed to methods and systems used to determine private set intersections (PSIs) and execute private database joins (PDJs). Some embodiments are characterized by binning techniques that enables PSI and PDJ methods to be performed by worker nodes in a computing cluster in parallel, thus reducing execution time. A first party computing system and a second party computing system can each tokenize their respective datasets, then assign the datasets to bins. The bins can each be padded with dummy tokens. Then the first party computing system and second party computing system can execute several parallel PSI on pairs of corresponding bins. The results can then be combined to produce a tokenized intersection set, which can then be detokenized to produce the set intersection.

[Continued on next page]

HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*