

Τεχνικές Εξόρυξης Δεδομένων 2008-2009

Άσκηση 2

Πολυχρονίου Ορέστης
1115200600176
Ψαλλίδας Φώτης
1115200600170

Αρχεία:

hac.c
Κώδικας των αλγορίθμων clustering

hac.h
Αρχείο επικεφαλίδας

main.c
Συνάρτηση main για το δεδομένο πρόβλημα

Makefile

Ορίσματα:

-i [input_file]
Εισαγωγή αρχείου εισόδου

-c [cluster_output_file]
Έξοδος σε αρχείο υπό μορφή clusters

-q [queue_output_file]
Έξοδος σε αρχείο υπό μορφή ουράς ενώσεων

-a [connection]
Επιλογή κριτηρίου ένωσης clusters
s για single
c για complete
a για average

-d [distance_metric]
Επιλογή μετρικής απόστασης
i για internal product
j για jaccard distance

Εκτελέσεις:

calirand.txt – single – product

```
>./hac -i calirand.txt -c out.txt -a s -d i
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.06 secs]
Clustering complete     [0.04 secs]
Result cluster file complete [0.00 secs]
```

calirand.txt – single – jaccard

```
>./hac -i calirand.txt -c out.txt -a s -d j
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.21 secs]
Clustering complete     [0.03 secs]
Result cluster file complete [0.00 secs]
```

calirand.txt – complete – product

```
>./hac -i calirand.txt -c out.txt -a c -d i
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.08 secs]
Clustering complete     [0.20 secs]
Result cluster file complete [0.00 secs]
```

calirand.txt – complete – jaccard

```
>./hac -i calirand.txt -c out.txt -a c -d j
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.25 secs]
Clustering complete     [0.34 secs]
Result cluster file complete [0.00 secs]
```

calirand.txt – average – product

```
>./hac -i calirand.txt -c out.txt -a a -d i
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.06 secs]
Clustering complete     [0.42 secs]
Result cluster file complete [0.00 secs]
```

calirand.txt – average – jaccard

```
>./hac -i calirand.txt -c out.txt -a a -d j
Dataset file loaded      [0.00 secs]
902 vectors - 1124 dimensions
Distances computed      [0.23 secs]
Clustering complete     [0.63 secs]
Result cluster file complete [0.00 secs]
```

titlesrand.txt – single – product

```
>./hac -i titlesrand.txt -c out.txt -a s -d i  
Dataset file loaded      [0.00 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [0.57 secs]  
Clustering complete     [0.29 secs]  
Result cluster file complete [0.00 secs]
```

titlesrand.txt – single – jaccard

```
>./hac -i titlesrand.txt -c out.txt -a s -d j  
Dataset file loaded      [0.00 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [1.63 secs]  
Clustering complete     [0.19 secs]  
Result cluster file complete [0.00 secs]
```

titlesrand.txt – complete – product

```
>./hac -i titlesrand.txt -c out.txt -a c -d i  
Dataset file loaded      [0.00 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [0.48 secs]  
Clustering complete     [1.84 secs]  
Result cluster file complete [0.00 secs]
```

titlesrand.txt – complete – jaccard

```
>./hac -i titlesrand.txt -c out.txt -a c -d j  
Dataset file loaded      [0.01 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [1.61 secs]  
Clustering complete     [1.73 secs]  
Result cluster file complete [0.00 secs]
```

titlesrand.txt – average – product

```
>./hac -i titlesrand.txt -c out.txt -a a -d i  
Dataset file loaded      [0.00 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [0.47 secs]  
Clustering complete     [3.53 secs]  
Result cluster file complete [0.00 secs]
```

titlesrand.txt – average – jaccard

```
>./hac -i titlesrand.txt -c out.txt -a a -d j  
Dataset file loaded      [0.01 secs]  
2223 vectors - 2348 dimensions  
Distances computed      [1.62 secs]  
Clustering complete     [3.52 secs]  
Result cluster file complete [0.01 secs]
```

Πολυπλοκότητα:

single linkage

Αρχικά βρίσκουμε για κάθε στοιχείο τον κοντινότερο του και τον κρατάμε σε ένα πίνακα [$O(n^2)$].

Έπειτα για κάθε στάδιο ένωσης επιλέγουμε την μικρότερη απόσταση από τον πίνακα των μικρότερων [$O(n)$] και ανανεώνουμε τον κοντινότερο για όλα τα άλλα clusters από τα 2 που ενώθηκαν. [$O(n)$].

Άρα συνολικά για c στάδια ενώσεων ($c < n$), έχουμε:
 $O(n^2 + (n+1)c) = O(n^2)$

complete linkage – average linkage

Αρχικά χτίζουμε για κάθε στοιχείο ένα σωρό, όπου κρατάμε στην κορυφή τον κοντινότερο [$O(n^2)$].

Έπειτα σε κάθε στάδιο επιλέγουμε από τις κορυφές τον μικρότερο και τον ενώνουμε, ενώ ανανεώνουμε τους υπόλοιπους σωρούς εξάγωντας τα στοιχεία των cluster που ενώθηκαν και εισάγωντας τις αποστάσεις από το νέο cluster [$O(\log n + n \log n)$]

Άρα συνολικά για c στάδια ενώσεων ($c < n$), έχουμε:

$$O(n^2 + c(n+1)\log n) = O(n^2 + cn \log n)$$

Αν πάρουμε την περίπτωση όπου θέλουμε να φτάσουμε μέχρι ένα cluster, τότε ισχύει $c = n-1$, άρα έχουμε:

$$O(n^2 + (n-1)n \log n) = O(n^2 \log n)$$

Οπτικοποιήσεις:

(Το νούμερο i εκφράζει το αντικείμενο στη σειρά i στο dataset.)

calirand.txt – single – product

C1: 195 602 306 452 749 822 249 479 740 809 821 789 723 722 538 603 803 617 462 52 127 707 584 663 762 586 574 276 14 69 847 804 720 502 374 373 268 213 104 103 8 35 403 783 402 375 75 288 546 633 567 64 182 871 877

C2: 259 174 447 611 892 209 255 369 414 475 316 317 319 341 446 653 4 203 181 296 511 514 527 536 576 639 725 286 162 150 118 83 84 241 111 191 625 645 180 90 141 661 665 812 896

C3: 315 876 870 800 765 652 620 21 41 851 745 550 540 274 9 190 724 766 901 897 895 878 863 860 857 855 853 852 846 842 836 831 826 817 792 787 786 774 763 760 755 751 747 744 735 733 728 719 718 714 704 696 690 683 678 675 674 667 660 656 655 640 623 613 612 609 601 597 594 589 587 585 566 556 554 549 529 515 510 509 507 503 493 492 489 463 426 421 418 404 379 361 312 301 283 277 275 270 265 261 245 240 236 232 231 227 226 223 222 210 207 205 201 200 185 178 177 172 171 168 165 164 155 154 146 144 132 125 122 107 106 101 99 97 94 92 88 85 82 80 77 76 72 68 67 61 39 37 36 32 30 28 23 22 20 7 5 531 153 13 680 869 782 634 659 780 781 894 742 677 703 900 885 867 864 813 784 692 682 592 290 238 143 29 60 100 220 192 53 730 668 624 411 291 866 43 108 199 267 593 741 49 299 380 419 451 460 485 600 628 830 27 342 456 520 551 562 413 406 405 356 334 491 629 849 51 271 423 708 89 139 140 395 396 731 157 608 779 761 644 134 287 865 824 297 561 432 647 126 501 215 262 648 596 357 457 464 518 522 595 614 672 694 230 883 345 796 829 775 579 578 577 547 445 344 2 280 881 119 544 828 166 560 477 257 131 187 214 621 55 242 662 816 152 91 6 771 739 158 81 128 525 526 669 516 530 412 138 366 715 790 727 605 116 300 688 778 793 186 129 16 606 31 10 0 15 115 508 558 607 427 428 573 429 70 244 250

C4: 376 534 519 548 351 450 545 498 279 320 513 298 401 430 487 494 512 893 861 838 837 681 564 506 499 484 483 482 471 453 420 394 392 391 390 389 388 387 386 348 347 304 449 448 385 384 383 382 349 313 850 666 441 438 437 381 365 332 331 330 329 328 327 326 321 318 407 398 310 371 370 309 397 467 743 393 408 434 874 658 42 78 239 542 543 845 700 563 359 50 773 772 159 202 148 686 188 676 458 615 630 631 217 670 147 235 619 679 650 627 646 711 734 767 695 691 637 167 73 71 12 59 795 801 814 825 237 524 472 523 832

C5: 497 638 465 459 377 588 833 618 535 466 363 362 333 439 368 338 335 746 444 337 311 308 322 339 340 436 372 433 570 673 840 899 495 818

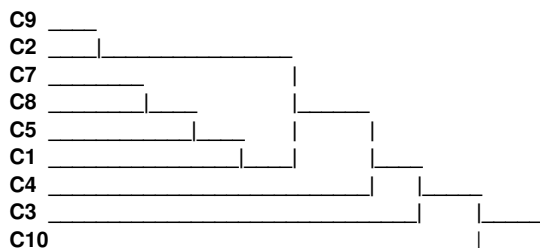
C6: 555 18 26 583 582 221 233 843 229 96 114 211 314 135 352 738 87 58 788 170 360 820 805 794 808 835 868 145 807 224 254 737 216 553 848 754 759 572 110 726 664 57 112 736 552 699 355 410 785 109 66 610 169 568 252

C7: 557 307 219 34 488 79 604 622 86 643 137 105 890 248 490 346 353 358 480 481 532 559 626 856

C8: 757 834 705 717 253 748 750 212 46 504 505 422 565 208 246 806 500 243 263 689 17 649 399 228 819 854 823 799 712 671 416 409 378 282 218 196 123 93 273 102 533 521 721 687 764 859 777 478 841 810 258 858 752 537 256 25 56 642 580 598 879 768 685 684 599 417 260 47 130 133 839 872 590 293 197 198 124 151 575 873 811 284 33 65 225 889 770 571 117 194 302 693 173 713 769 776 862 657 636 496 569 716 791 189 882 272 343 654 701 702 176 193 292 802 641 651 541 98 517 528 706 880 815 206 156 63 121 884 797 179 38 120 11 474 758 473 886 269 285 756 149 698 616 469 470 539 303 431 247 278 251 44 697 753 424 425 400 591 364 710 281 891 266 888 264 461 234 635 204 289 160 142 113 183 74 163 62 581 48 95 45 732 40 367 19 709 3 729 887 1 161

C9: 798 175 454 455 435 440 336 476 468 415 350 354 323 324 325 442 443 827 136 632 294 295

C10: 898 875 844 305 184 24 54 486



calirand.txt – single – jaccard

C1: 125

C2: 153 13 680 531

C3: 199

C4: 215 262 647 432 727 126 501 715 648 605 116 300 0 15 10 31 16 129 186 115 646 606 778 665 688 250 429 573 70 244
427 428 607 110 572 558 508 661 790 793 142 160 167 625 162 639 725 118 83 84 286 241 111 191 190 414 461 264 45 732
5 787 728 566 489 223 210 144 107 61 28 30 3 729 7 19 709 22 23 32 36 37 40 367 51 72 74 76 78 82
85 92 97 99 109 122 132 134 287 154 155 164 165 169 176 193 292 177 185 200 201 204 289 205 225 889 770 226 227 231
232 234 635 236 254 258 810 478 841 25 858 56 256 752 537 777 859 764 687 721 261 265 275 283 284 33 65 502 312 315
876 21 620 652 41 765 800 242 695 361 364 379 400 591 418 421 423 424 425 426 463 492 493 497 503 507 515 527 536 539
303 431 554 556 575 585 587 601 612 616 469 470 623 640 655 656 667 674 675 678 683 706 711 714 719 735 736 744 755 760
763 786 792 826 842 852 853 855 863 871 14 574 69 847 374 373 268 213 103 8 35 276 586 762 804 720 878 879 768 880
884 797 179 38 120 897 58 788 170 360 151 156 163 197 198 740 809 249 479 749 822 20 44 697 753 48 95 49 741 593
68 77 80 101 114 229 211 119 544 146 166 828 824 297 865 779 171 175 798 178 189 716 569 882 207 222 224 240 243 245
247 278 251 252 568 253 834 266 888 849 380 269 285 886 270 271 277 281 891 293 590 299 509 510 529 542 543 50 359 159
202 773 563 772 298 496 636 862 657 571 580 598 584 663 594 597 621 55 622 624 291 43 108 866 187 257 131 477 214 47
417 685 411 560 599 684 730 267 608 637 767 217 71 12 248 559 490 856 480 626 481 358 346 353 532 59 105 890 658 660
662 816 668 691 696 699 355 552 18 555 583 582 221 233 410 26 702 708 710 718 737 216 553 754 848 747 751 757 774 795
430 401 376 519 534 548 545 827 443 442 325 323 324 476 468 415 370 371 372 322 339 340 436 279 498 333 439 368 335 338
337 444 311 746 336 351 354 350 861 837 681 484 483 482 471 453 420 392 391 390 389 388 387 386 347 348 394 499 893 308
304 363 618 362 466 433 570 535 833 459 465 377 588 838 840 673 450 357 832 518 614 464 487 513 320 595 629 672 801 814
334 356 406 405 342 413 435 440 402 546 633 562 628 451 485 551 456 520 460 600 27 830 341 511 319 316 317 514 576 457
491 522 783 506 564 638 743 419 454 455 494 512 73 75 288 375 475 642 785 870 899 90 180 150 91 776 6 769 771 739
158 81 128 255 209 713 369 403 528 596 611 892 174 447 896 734 802 343 272 654 701 808 835 868 818 836 846 857 860 873
877 42 53 66 610 86 87 738 98 517 651 541 641 117 124 127 803 617 52 462 821 789 723 722 538 603 602 130 192 260
263 446 500 524 472 523 237 504 505 561 645 669 516 530 707 412 138 366 692 694 761 811 843 845 867 881 1 161 887 11
474 473 758 39 67 89 94 100 220 106 133 135 141 148 172 181 203 194 302 693 173 195 239 259 273 93 854 799 712 671
282 123 196 102 409 416 104 218 378 399 533 819 228 823 296 306 452 352 404 521 549 565 815 206 63 121 183 208 589 609
653 676 188 686 458 615 630 631 670 147 235 627 393 872 874 650 467 365 850 666 438 437 381 332 331 330 329 328 327 326
318 321 449 448 385 384 383 382 313 349 407 310 398 309 408 441 434 397 679 495 619 700 704 733 780 781 894 60 885 813
682 592 290 29 143 238 246 864 659 869 634 782 900 677 703 742 784 731 157 745 9 851 540 550 274 395 396 812 817 820
805 794 825 831 839 901 4 46 113 212 644 689 17 649 726 791 806 748 750 422 705 717 139 140 844 305 486 898 24 54
184 875 64 182 567 88 96 112 759 57 664 525 526 152 34 488 557 219 307 698 149 756 137 643 604 79 145 807 168 230
883 345 796 829 579 578 577 547 445 344 2 280 775 301 314

C5: 581 62

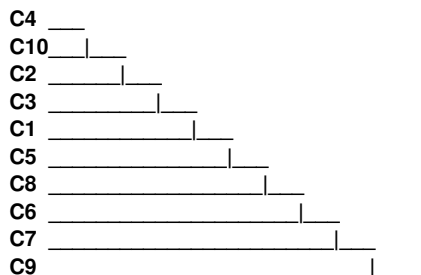
C6: 613

C7: 632 294 295 136

C8: 690

C9: 724 766

C10: 895



calirand.txt – complete – product

C1: 4 203 880 865 824 297 561 818 231 667 113 183 765 41 652 800 21 620 30 886 269 285 119 544 219 557 307 34 488 99
690 234 635 860 359 563 50 542 543 49 741 190 274 851 745 550 9 540 641 98 517 541 651 80 315 876 130 133 473 758
879 811 596 761 206 815 121 63 156 797 884 120 38 179 76 495 899 872 235 147 619 51 271 679 181 314 812 366 138 715
88 312 169 458 615 67 774 418 266 888 126 647 124 151 497 153 531 74 163 241 191 111 286 77 554 176 702 96 731 658
675

C2: 11 474 491 319 296 511 435 440 336 476 342 320 513 446 653 375 75 288 475 317 576 403 783 341 316 514 609 248 12 559
856 105 890 68 400 591 145 807 802 817 132 301 718 644 134 287 82 594 173 693 199 593 91 152 714 704 175 798 454 455
19 709 423 823 193 292 101 477 187 131 257 730 411 291 624 45 732 689 237 524 472 523 504 505 825 845 90 141 645 180
150 298 79 604 174 447 896 259 611 892 17 649 665 683 110 572 465 459 638 466 673 444 337 746 833 338 333 439 368 335
308 311 433 570 340 436 372 322 339 840 362 535 363 618 377 588 26 584 663 877 707 762 57 664 719 230 694 756 149 698
86 622 171 587 20 878 201 492 107 839 165 820 794 805 109 751 204 289 97 493 239 431 303 539

C3: 14 276 574 586 22 882 195 616 469 470 53 135 197 198 293 590 735 228 819 144 424 425 585 7 744 502 284 33 65 48
95 724 766 770 556 251 792 42 78 240 515 421 261 270 232 275 216 553 157 608 416 93 273 854 196 671 123 602 102 282
799 378 712

C4: 54 184 898 875 844 24 305

C5: 59 70 573 427 428 217 630 631 244 429 767 695 637 167 71 73 18 555 496 636 657 862 245 674 142 160 763 223 871 55
621 208 565 580 598 148 343 272 654 701 793 527 536 85 786 172 640 58 788 868 808 835 23 863 507 601 646 508 558 705
748 750 426 409 533 379 264 461 849 642 402 546 633 380 419 39 887 1 161 612 246 677 703 238 592 900 869 782 634 659
60 290 894 780 781 29 143 885 864 813 784 682 742 13 680 254 224 726 399 218 710 510 252 568 678 250 613 711 125 412
530 787 222 566 117 571 194 302 364 265 267 87 738 170 360 61 352 754 848 209 648 6 255 776 713 769 66 610 490 358
627 650 481 353 626 480 346 532 467 434 441 850 437 370 666 438 381 365 332 331 330 329 328 327 326 318 321

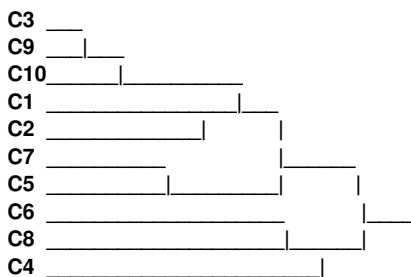
C6: 486 396 395 779

C7: 687 764 108 43 866 521 721 178 597 146 708 89 139 140 37 852 192 867 757 834 207 422 575 516 669 855 106 306 452 116
605 790 115 31 300 10 0 15 28 552 410 355 699 791 189 716 795 463 258 810 478 841 36 462 617 52 127 803 821 789
723 722 538 603 40 367 897 589 247 278 100 220 188 676 686 700 655 509 253 717 853 736 785 136 632 294 295 826 205 283
215 262 185 414 639 725 56 537 25 256 858 752 777 859 846 210 281 891 92 361 168 501 94 895 166 560

C8: 734 506 838 893 861 420 390 386 681 482 393 874 449 397 407 398 310 371 448 385 384 383 382 313 349 499 837 443 827 415
468 442 325 323 324 484 408 309 743 453 391 387 347 564 483 394 389 350 354 471 392 388 348 304 548 351 450

C9: 760 225 889 755 226 529 759 623 155 733 62 581 369 691 706 857 277 829 2 547 883 345 796 445 578 775 579 344 280 577
249 740 809 479 749 822 118 83 84 211 221 582 233 583 32 836 430 401 512 376 519 534 498 279 545 72 549 177 661 162
625 464 551 456 520 460 628 485 451 600 487 494 413 406 405 356 334 562 457 522 814 357 518 832 801 672 629 595 614

C10: 828 881 831 696 236 607 688 778 186 129 16 606 44 697 753 569 528 27 830 46 299 212 668 154 806 656 242 662 816 5
901 747 227 404 112 525 526 842 214 599 417 684 873 768 685 47 260 489 670 870 122 503 182 64 567 660 164 3 729 728
229 114 843 432 727 200 737 243 263 500 692 137 643 847 771 739 158 81 128 213 720 773 202 159 772 104 35 69 103 8
268 804 373 374



calirand.txt – complete – jaccard

C1: 89 708 139 140

C2: 96 843 211 114 229

C3: 136 632 294 295

C4: 199 593 605

C5: 224 254 57 664 726 110 572

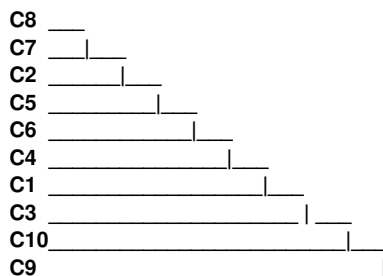
C6: 239 899 638 459 465 308 433 570 340 436 372 322 339 377 588 840 833 535 363 466 362 618 446 653 296 403 783 633 402 546
512 401 430 375 75 288 475 73 514 576 341 319 317 316 511 642 419 600 460 551 628 451 485 456 520 673 746 311 337 444
494 487 513 336 468 476 320 279 498 342 435 440 333 439 368 335 338 562 413 405 334 356 406 870 522 27 830 457 464 595
801 629 672 814 357 832 518 614

C7: 314 679 181 203 791 189 716 141 645 90 180 175 798 454 455 298 376 519 534 42 658 230 694 66 169 610 79 604 86 622
137 643 105 890 856 490 12 248 559 613 524 472 523 825 237 504 505 845 700 563 50 359 542 543 772 773 159 202 720 104
804 502 69 213 373 847 374 103 35 8 268 46 212 299 380 849 724 766 302 693 194 117 713 641 541 98 517 651 145 807
410 736 785 699 355 552 584 663 877 26 707 259 611 892 896 174 447 53 192 267 867 668 560 730 411 291 624 692 100 220
869 634 782 659 780 781 894 677 703 784 900 742 246 238 864 60 29 143 885 813 682 290 592 561 865 297 824 596 811 779
274 395 396 130 133 619 495 872 147 235 260 608 684 599 685 214 47 417 4 710 702 802 500 806 689 243 263 839 13 680
690 521 124 873 151 575 293 590 197 198 284 33 65 412 138 366 881 119 544 153 531 34 488 557 219 307 369 648 209 255
91 776 6 769 771 739 158 81 128 62 581 670 630 631 167 59 637 217 71 767 427 428 691 70 573 244 429 157 731 208
565 688 665 778 116 300 115 646 508 558 607 16 129 186 606 0 15 10 31 187 257 131 477 537 752 25 858 56 256 478
841 777 859 43 108 866 721 687 764 173 352 17 649 190 745 9 851 540 550 414 491 286 241 111 191 135 528 271 253 834
113 183 125 717 705 748 750 895 306 452 749 822 249 479 740 809 11 474 567 64 182 156 880 815 121 63 206 74 163 625
661 162 150 639 725 195 215 262 647 432 727 272 343 654 701 112 759 152 662 816 794 805 820 58 788 808 695 835 868 168
228 819 218 399 292 823 102 273 602 378 416 93 123 282 854 799 712 196 671 793 527 536 571 770 715 126 501 301 88 473
758 264 461 901 589 1 161 887 250 49 741 106 882 831 704 67 127 52 617 462 803 821 789 723 722 538 603 676 188 686
172 549 609 817 580 598 790 525 526 148 55 621 242 800 765 41 21 620 652 754 848 733 669 516 530 657 496 636 862 404
757 775 883 345 796 445 578 2 547 829 579 344 280 577 812 142 160 94 87 738 39 509 860 178 277 176 193 222 281 891
245 569 19 709 510 146 270 118 83 84 221 582 233 583 207 409 533 45 732 44 697 753 251 258 810 747 846 737 216 553
240 857 170 360 171 68 594 458 615 718 77 529 101 269 285 886 774 315 876 751 80 266 888 252 568 698 149 756 818 166
828 768 879 836 225 889 247 278 696 18 555 795 660 597 616 469 470 14 574 762 276 586 48 95 20 507 361 82 283 109
853 792 755 735 364 165 76 554 492 177 154 601 871 23 261 878 786 744 22 3 729 122 852 797 884 120 38 179 463 706
423 683 842 760 92 585 503 719 855 422 678 227 40 367 431 303 539 234 635 826 379 667 226 640 201 312 99 36 32 7
515 497 714 426 85 72 275 231 655 204 289 200 623 400 591 493 711 424 425 421 675 418 674 51 265 97 587 236 164 612
863 763 185 155 37 132 232 656 205 78 556 897 5 61 728 30 144 28 566 489 223 210 107 787

C8: 644 761 134 287

C9: 734 450 351 545 548 397 309 408 564 506 838 304 893 499 394 350 482 861 483 453 837 484 471 420 392 391 390 389 388 387
386 348 347 681 874 650 393 627 532 480 481 358 353 346 626 743 370 415 354 443 827 442 325 323 324 434 467 371 407 310
398 441 449 448 385 384 383 382 313 349 365 850 437 666 438 381 332 331 330 329 328 327 326 318 321

C10: 875 54 184 24 305 844 486 898



calirand.txt – average – product

C1: 40 367 424 425 23 245 616 469 470 463 423 11 474 473 758 39 655 19 709 658 42 78

C2: 44 697 753 612 601 594 818 264 461 863 855 510 507 497 421 404 277 270 261 236 231 226 623

C3: 48 95 252 568 169 109 66 610 895 860 182 64 567 204 289 62 581 795 45 732

C4: 120 797 884 38 179 897 878 826 792 786 831 887 1 161 251 247 278 3 729 886 269 285

C5: 445 579 775 280 577 829 344 578 2 547 883 345 796 234 635 74 163 774 719 714 696 683 675 667 656 613 609 597 587 817

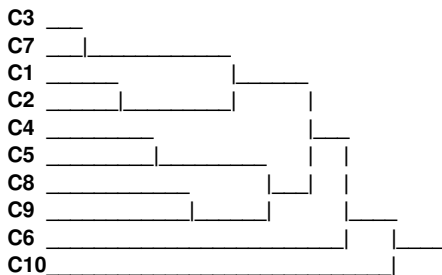
C6: 496 636 657 862 882 569 791 189 716 86 622 79 604 137 643 756 149 698 219 557 307 34 488 170 360 87 738 215 262 432
727 501 126 647 669 516 530 138 715 366 412 848 737 754 216 553 110 572 145 807 254 224 726 142 160 49 741 199 593 790
525 526 57 664 112 759 458 615 630 631 767 167 637 217 59 71 244 429 70 573 427 428 250 661 665 688 778 115 646 605
116 300 0 15 10 31 793 129 186 508 558 607 16 606 524 472 523 237 825 504 505 306 452 770 225 889 249 749 822 479
740 809 52 127 462 617 803 789 538 603 821 722 723 644 611 892 896 259 174 447 119 544 828 881 134 287 865 824 297 761
364 710 399 195 823 193 292 409 533 228 819 416 196 671 712 218 104 282 799 102 93 273 854 378 123 602 686 188 676 148
173 693 194 302 713 117 571 820 794 805 808 835 868 58 788 369 691 527 536 91 152 769 209 648 776 6 255 845 50 563
359 700 542 543 502 213 158 739 771 81 128 374 103 720 847 35 69 373 8 268 804 159 772 202 773 136 632 294 295 4
203 181 314 229 114 843 352 53 135 13 680 153 531 706 641 98 517 541 651 113 183 880 206 815 121 63 156 802 176 702
272 343 654 701 689 17 649 662 816 596 267 192 867 100 220 243 263 500 692 900 29 143 784 60 290 813 238 592 864 246
682 885 742 677 703 894 780 781 634 659 782 869 258 521 721 687 764 858 56 537 25 256 752 777 859 810 478 841 166 560
291 624 108 43 866 411 668 730 477 131 257 208 565 851 190 274 9 540 550 745 879 839 575 873 124 151 293 590 284 561
33 65 197 198 395 779 396 811 708 89 139 140 96 731 157 608 187 214 417 599 685 768 684 47 260 175 798 239 757 834
51 271 253 717 266 888 642 849 212 46 299 422 806 705 748 750 871 574 586 762 14 276 736 410 785 118 83 84 211 221
582 233 583 18 555 552 355 699 584 663 877 26 707 180 90 141 645 150 298 111 286 191 241 162 625 414 639 725 55 621
670 870 315 876 800 41 652 242 695 765 21 620 459 638 465 495 580 598 130 133 235 679 872 147 619 230 694 856 105 890
490 248 12 559 376 430 401 512 475 317 73 511 446 653 319 296 514 341 316 576 75 288 375 402 403 783 546 633 734 480
353 626 481 358 346 532 528 27 830 672 357 518 832 801 487 494 464 595 614 629 457 522 814 454 455 491 406 334 562 413
356 405 551 456 520 336 342 435 440 840 466 362 535 377 588 833 363 618 673 433 570 899 368 333 439 335 338 308 311 444
337 746 322 339 372 340 436 393 874 627 650 519 534 548 351 450 320 513 498 279 545 893 838 506 484 482 453 394 391 389
387 348 350 304 564 861 837 499 483 471 420 392 390 388 386 347 681 380 419 460 628 485 451 600 309 408 434 467 449 382
383 384 385 448 313 349 443 323 324 325 354 442 827 476 415 468 743 397 326 327 332 666 328 329 437 850 330 331 438 441
365 381 318 321 370 310 371 398 407

C7: 589 566 554 515 763 585 301 222 207 201 185 177 171 165 155 146 76 68 61 36 30 22 7 852

C8: 755 747 735 728 556 549 509 503 492 529 857 853 836 846 812 281 891 704 690 678 674 660 640 361 210 205 200 178 172 168
164 132 122 106 99 94 88 82 77 72 724 766

C9: 842 787 760 751 744 733 718 711 493 431 303 539 400 591 489 426 418 379 312 283 275 265 240 232 227 223 154 144 125 107
101 97 92 85 80 67 37 32 28 20 5 901

C10: 875 486 24 305 898 844 54 184



calirand.txt – average – jaccard

C1: 4

C2: 46 212 296 446 653 145 807 224 726 89 708 139 140 690 125 521 839 710 13 680 153 531 113 183 190 745 79 604 137 643
689 17 649 239 899 641 541 98 517 651 845 825 504 505 237 524 472 523 895 301 168 88 195 163 625 53 192 596 230 314
181 203 500 263 243 692 62 581 208 565 135 528 173 352 91 152 369 209 648 302 693 194 713 117 571 141 645 90 180 259
896 174 447 611 892 11 474 306 452 42 658 96 843 211 114 229 414 162 150 639 725 761 561 779 865 297 824 57 664 112
759 567 64 182 86 622 757 271 253 834 100 220 901 831 817 812 733 704 609 589 549 404 172 106 94 67 39 87 738 133
148 267 34 488 557 219 307 250 790 525 526 638 175 798 454 455 494 403 783 475 73 512 401 430 375 75 288 633 402 546
576 514 341 316 511 317 319 736 785 456 520 419 460 600 673 308 311 746 377 588 840 459 465 833 535 363 466 362 618 337
444 333 439 368 335 338 433 570 340 436 372 322 339 27 830 522 457 320 487 513 342 435 440 562 413 405 334 356 406 814
801 672 629 595 464 357 832 518 614 743 351 450 376 279 498 519 534 545 548 564 506 434 467 650 874 393 627 532 490 480
353 358 481 346 626 838 304 336 468 476 397 371 407 310 398 354 551 628 451 485 309 408 893 499 394 350 484 453 387 388
861 482 391 392 837 471 389 390 347 681 483 420 348 386 441 449 382 383 384 385 448 313 349 370 415 442 443 323 324 325
827 365 332 666 328 329 381 850 437 438 330 331 318 321 326 327 266 888 299 642 380 849 215 262 458 615 630 631 691 59
427 428 70 573 244 429 637 217 71 767 157 731 702 802 272 343 654 701 49 741 593 116 605 793 255 527 536 293 590 811
873 197 198 33 65 284 502 806 717 705 748 750 881 119 544 260 130 608 684 599 768 879 685 214 47 417 794 805 820 58
788 808 835 868 580 598 679 619 495 872 147 235 167 661 688 665 778 115 646 508 558 606 607 16 129 186 300 0 15 10
31 882 569 189 716 473 758 142 160 55 621 662 816 584 663 877 410 699 355 552 26 707 18 555 118 83 84 221 582 233
583 694 105 890 856 12 248 559 687 721 668 730 560 411 291 624 43 108 866 478 841 764 777 859 258 810 537 187 257 131
477 752 25 858 56 256 124 151 575 647 432 727 749 822 249 479 740 809 412 138 366 715 126 501 156 880 815 121 63 206
867 784 659 677 703 869 634 782 900 742 780 781 894 864 246 238 60 885 29 143 290 592 682 813 169 610 252 568 251 247
278 44 697 753 281 891 166 828 110 572 860 857 846 836 818 795 774 751 747 718 696 660 597 594 529 510 509 277 270 245
240 222 207 178 171 146 101 80 77 68 20 533 670 870 315 876 695 242 800 765 41 21 620 652 700 676 188 686 298 563
50 359 542 543 669 516 530 1 161 887 264 461 657 496 636 862 602 228 819 218 399 273 823 378 104 416 409 102 93 854
123 282 196 671 712 799 762 586 276 14 574 754 848 737 216 553 127 52 617 462 803 789 821 538 603 722 723 772 773 720
776 6 769 158 739 771 81 128 159 202 804 69 8 268 35 373 103 374 213 847 770 225 889 616 469 470 491 286 241 111
191 170 360 134 287 48 95 45 732 19 709 775 883 345 796 2 547 344 578 445 579 829 280 577 698 149 756 269 285 886
292 176 193 274 395 396 9 851 540 550 120 797 884 38 179 431 303 539 424 425 400 591 234 635 204 289 40 367 3 729
897 878 871 863 855 853 852 842 826 792 786 763 760 755 744 735 719 714 711 706 683 678 675 674 667 656 655 640 623 612
601 587 585 556 554 515 507 503 497 493 492 463 426 423 422 421 418 379 364 361 312 283 275 265 261 254 236 232 231 227
226 205 201 200 185 177 165 164 155 154 132 122 109 99 97 92 85 82 78 76 74 72 51 37 36 32 23 22 7 5
107 787 210 28 566 223 30 144 489 61 728

C3: 66

C4: 136 632 294 295

C5: 199

C6: 613

C7: 644

C8: 724 766

C9: 791

C10: 875 734 184 54 24 898 486 305 844

