drift from initial within g 1.0 sample_id similarity 8.0 baseline empirical cosine s (1,1) (1,2) (1,3) (1,4) (1,5) (1,6) repetition pair