Methods overlap for CM pairs in test_data,
total #CM pairs=658,
similarity threshold for uniqueness = 0