**Protocol Description**

Alice and Bob are supposed to do a secure two-party record linkage with the following protocol, which consists of the initialization state and linkage stage.

*Initialization:*

1. Alice
    a. generates the Paillier key pair (PrivK, PubK)
    b. sends PubK to Bob
    c. encodes each of her records into a Bloom Filter (e.g., see the method in [1])
    d. sends the parameters used to generate Bloom Filters to Bob

2. Bob
    a. encodes each of his records into a Bloom Filter in the same parameters as Alice

*Linkage:*

Let BLa be a Bloom Filter of a record from Alice, and let BLb be a Bloom Filter of a record from Bob. That is, BLa and BLb are two binary arrays, and suppose their length is L.

The following steps calculate the similarity ratio between BLa and BLb.

1. Alice
    a. for i=1 to L, generates BLa'[i] = Enc(PubK, BLa[i])
    b. sends the encrypted bloom filter BLa' to Bob

2. Bob
    a. let M be the modulus used by the Paillier encryption
    b. generates three random numbers r, e' and e0', ensuring that 2*r*L+e' < M, r*L+e0' < M, e'/r < 0.001, and e0'/r < 0.001
        i. the conditions e'/r < 0.001 and e0'/r < 0.001 should be set as parameter, which can be easily adjusted for different accuracy
    c. calculates homomorphically H = $(\sum_{i=1}^{L} BLb[i] * BLa'[i])$*r + Enc(PubK, e0') and homomorphically N = $(\sum_{i=1}^{L} BLa'[i]$+ Enc(PubK, $\sum_{i=1}^{L} b[i]))$*r + Enc(PubK, e')
        i. For convenience the above description assumes the operations (* and +) are the same in the plaintext and ciphertext domains; for Paillier system, they need to be replaced accordingly (+ by *, * by ^).
    d. sends H and N to Alice
3. Alice
    a. decrypts h = Dec(PrivK, H) and n = Dec(PrivK, N)
    b. calculates the approximate Dice coefficient as 2*h/n

Reference

[1] Rainer Schnell, Tobias Bachteler and Jörg Reiher. "Privacy-preserving record linkage using Bloom filters", BMC Medical Informatics and Decision Making, 2009.