# Strong Tie Prediction (STP) Candidate Source

Provides accounts with a high probability of potential mutual follows between the target user and the candidates.

## STP: Strong Tie Prediction

STP refers to the prediction of p(MutualFollow) for a given pair of users, which powers the concept of People You May Know (PYMK).

For training, positives are existing mutual follows and negatives are mutual follows of your follows. Features help distinguish between friends and friends-of-friends.

For inference, candidates are the topK mutuals of your follows. These are rescored, and we only send the topN to the product or next re-ranker.

### Online STP

Online STP generates a pool of candidates who are then ranked via a lightweight ranker.

It does this through a two-hop expansion of the mutual follow graph of users, where the first-degree neighbor is another user who has a link with the target user from following types:

\* Mutual Follow

\* Outbound phone contacts

\* Outbound email contacts

\* Inbound phone contacts

\* Inbound email contacts

The second-degree neighbor can only be a mutual follow link.

Currently, online STP can only perform the two-hop expansions on new users (<= 30 days since signup) due to compute resource constraints.

Features used for the lightweight ranker:

\* realGraphWeight: real graph weight between user and first degree nodes

\* isForwardEmail: whether the candidate is in the user's email book

\* isReverseEmail: whether the user is in the candidate's email book

\* isForwardPhonebook: whether the candidate is in the user's phone book

\* isReversePhonebook: whether the user is in the candidate's phone book

\* numMutualFollowPath: number of mutual follow path between the user and the candidate

\* numLowTweepcredFollowPath: number of mutual low TweepCred path between the user and the candidate

\* Tweepcred is a social network analysis tool that calculates the influence of Twitter users based on their interactions with other users. The tool uses the PageRank algorithm to rank users based on their influence.

\* hasForwardEmailPath: is there a third user x in the user's email book that connect user -> x -> candidate?

\* hasReverseEmailPath: is there a third user x in the user's reverse email book that connect user -> x -> candidate?

\* hasForwardPhonebookPath: is there a third user x in the user's phonebook that connect user -> x -> candidate?

\* hasReversePhonebookPath: is there a third user x in the user's reverse phonebook that connect user -> x -> candidate?

### Offline STP

Offline STP is powered by Pointwise Mutual Information, which measures the association between two users based on Twitter's mutual follow graph.

An offline job generates candidates based on the overlap between their Mutual and Addressbook Follows and that of the target user. Candidates are then made available online.

Candidates in OfflineSTP are "accounts that have a high overlap of mutually-followed accounts with an account in your follow graph."

This can potentially mean that OfflineSTP has a bigger reach than OnlineSTP.

For example, in the provided diagram, B and C have a high overlap of mutual follows, so it would be considered a candidate for A that is three hops away.

![img.png](img.png)

Overall, STP is a useful candidate source for generating potential follow recommendations based on strong ties between users, but it should be used in conjunction with other candidate sources and re-rankers to provide a well-rounded set of recommendations for the target user.