Empirical Analysis of Online Social Networks

Yu Chen

Dataset Overview
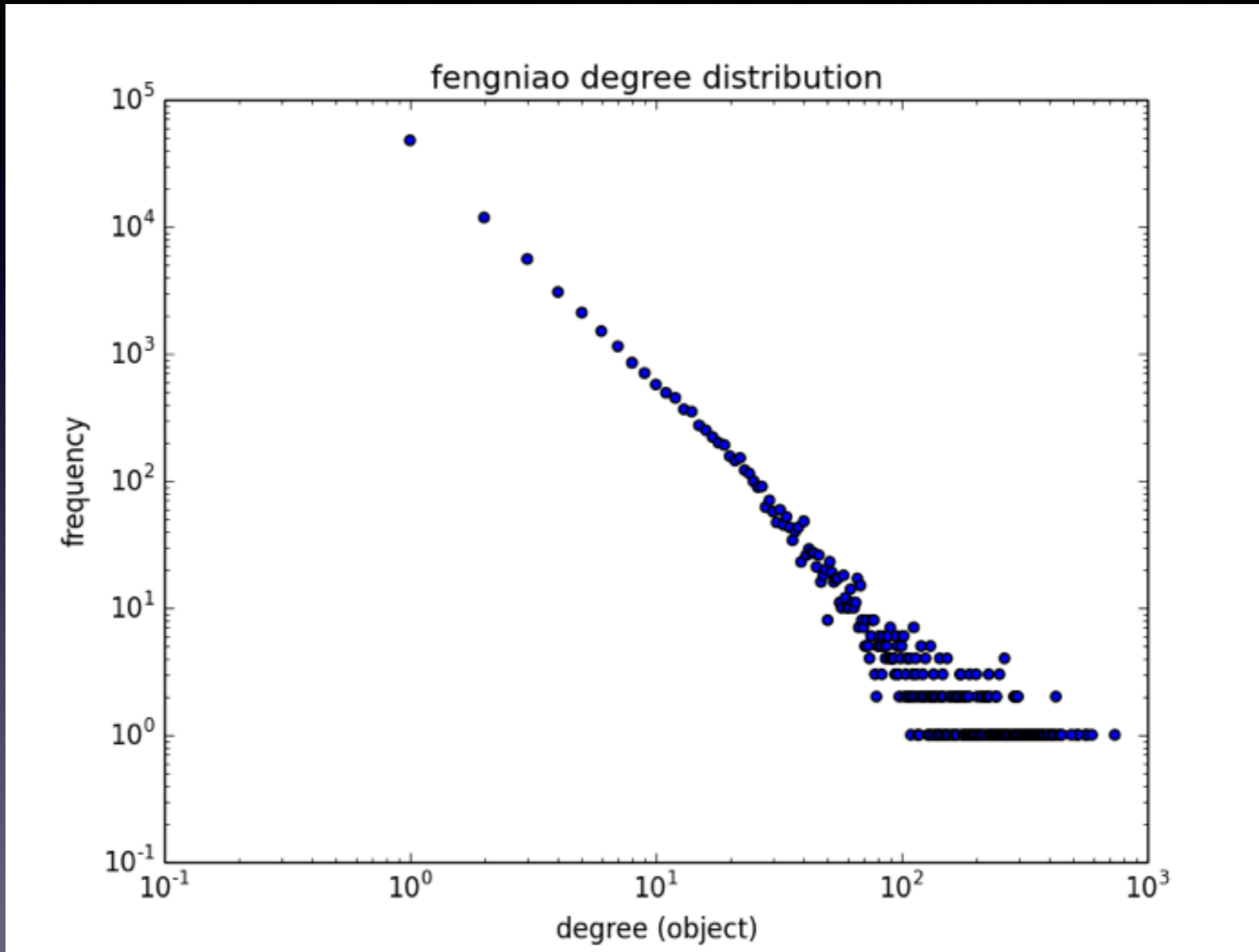
Dataset	Users	Pages (Objects)	Records
Fengniao (forum)	10, 000	80, 148	322, 399

Degree Distribution (User)

fengniao degree distribution
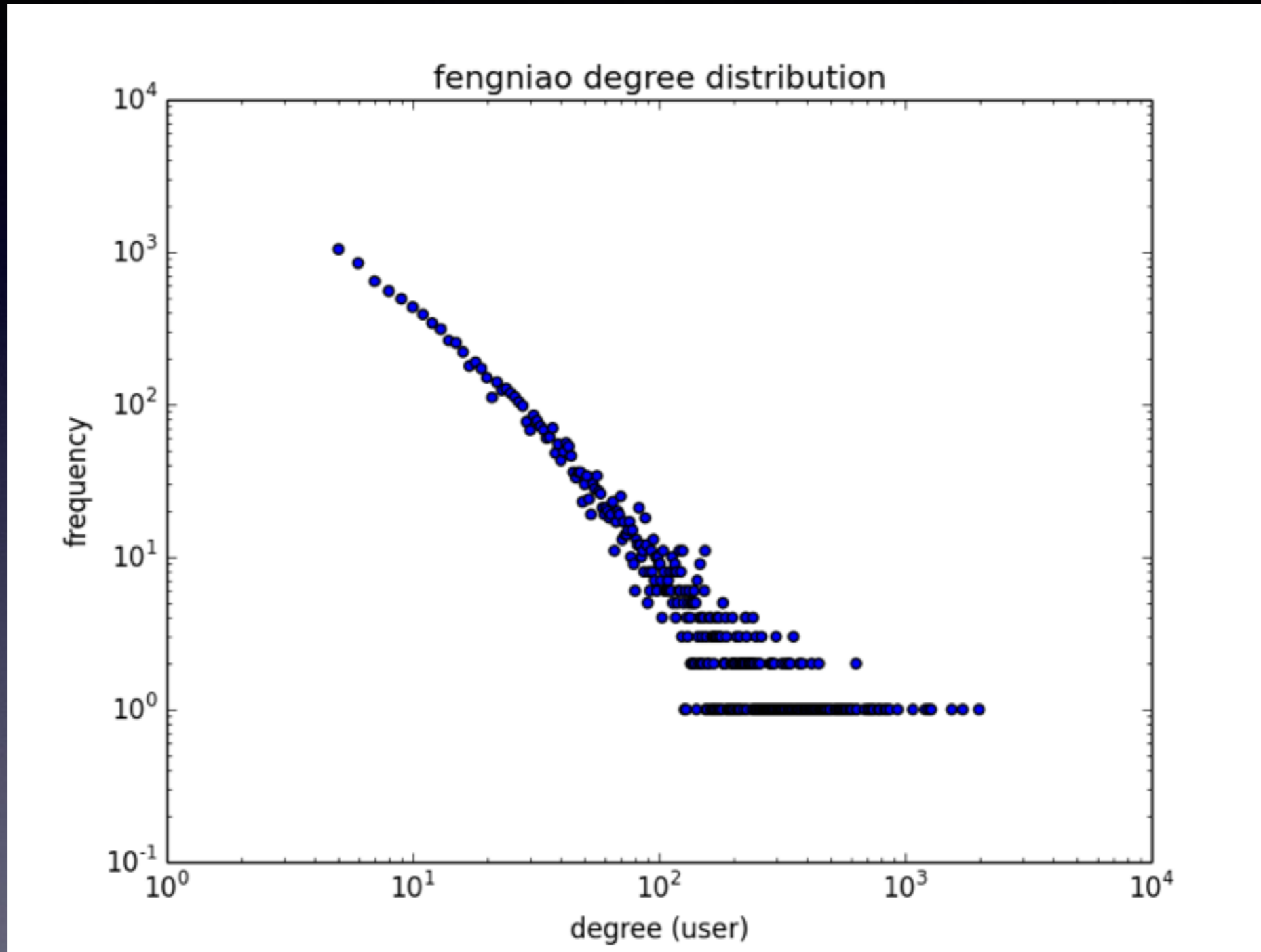
Follows the power law distribution

# Degree Distribution (Object)

Follows the power law distribution

Collaborative Similarity

The collaborative similarity[1] of a user is defined as the average similarity among the user's selected objects. It indicates how a user' interests are concentrated or dispersed.

The collaborative similarity[1] of an object is defined as the average similarity among the users who select the object. It indicates how similar the users who select the same object are.
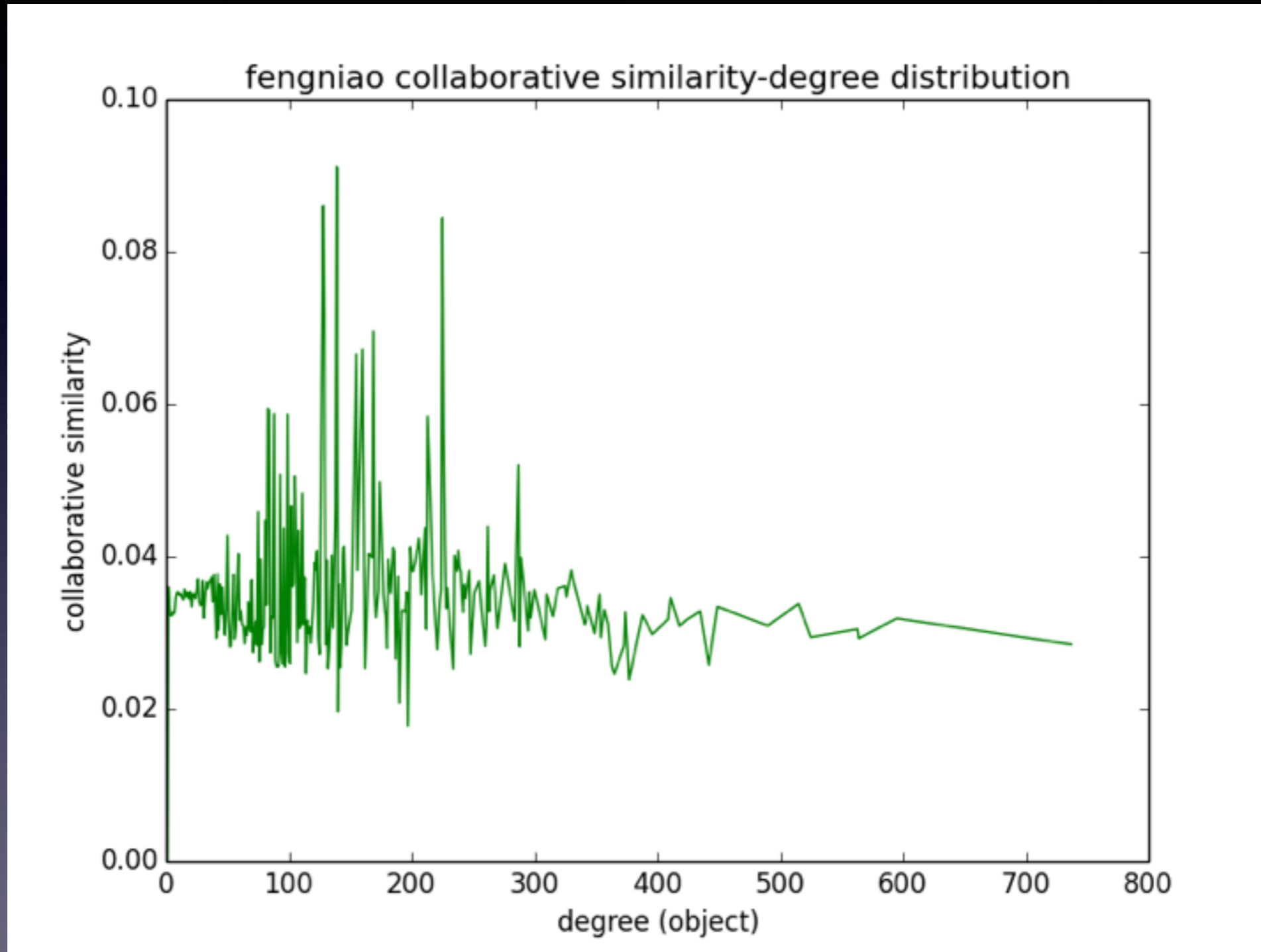
When we combine the above measurement with degree, we are going to find some interesting human behavioral patterns which we will see soon.

Collaborative Similarity (User)
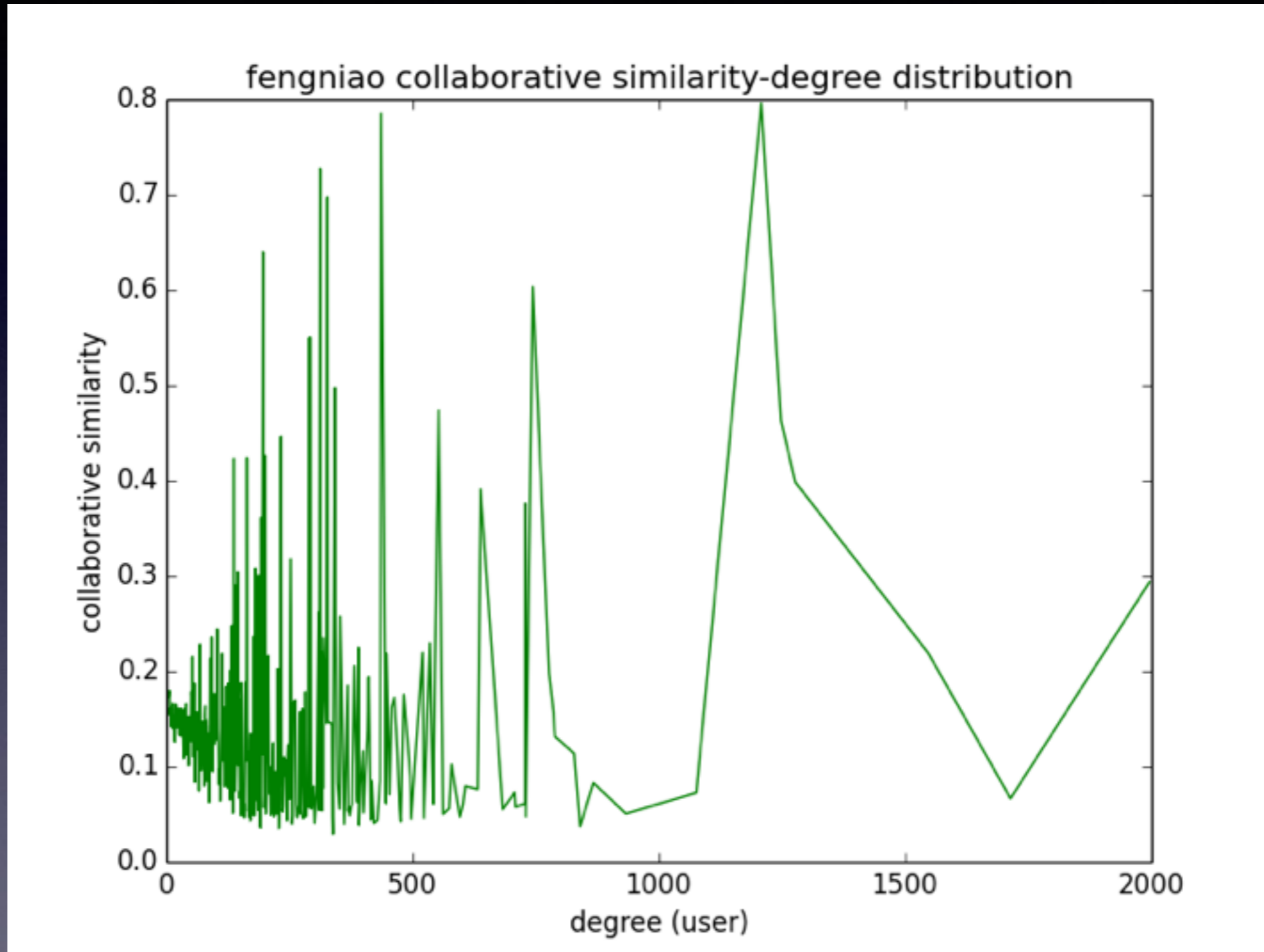
Fresh (small degree) users tend to have more concentrated interests.[1]

Collaborative Similarity (Object)

Users who select the same unpopular (small degree) object would be more similar than those who select the same popular (large degree) object.[1]

Nearest Neighbors' Degree

The nearest neighbors' degree for a **user** is defined as the average degree of all the objects connected to the user.
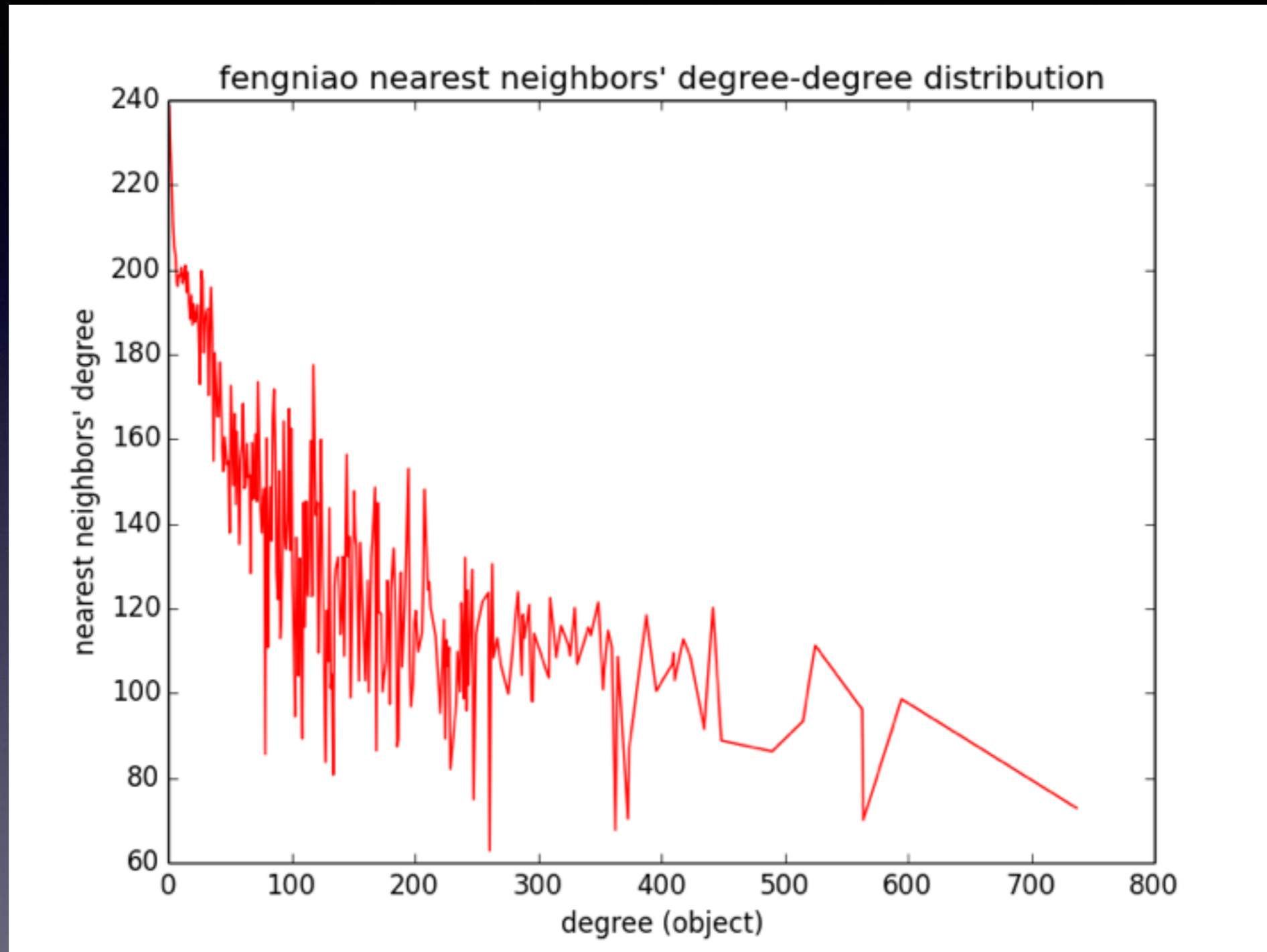
The nearest neighbors' degree for an **object** is defined as the average degree of all the users who select the object.

Nearest Neighbors' Degree (User)

fengniao nearest neighbors' degree-degree distribution
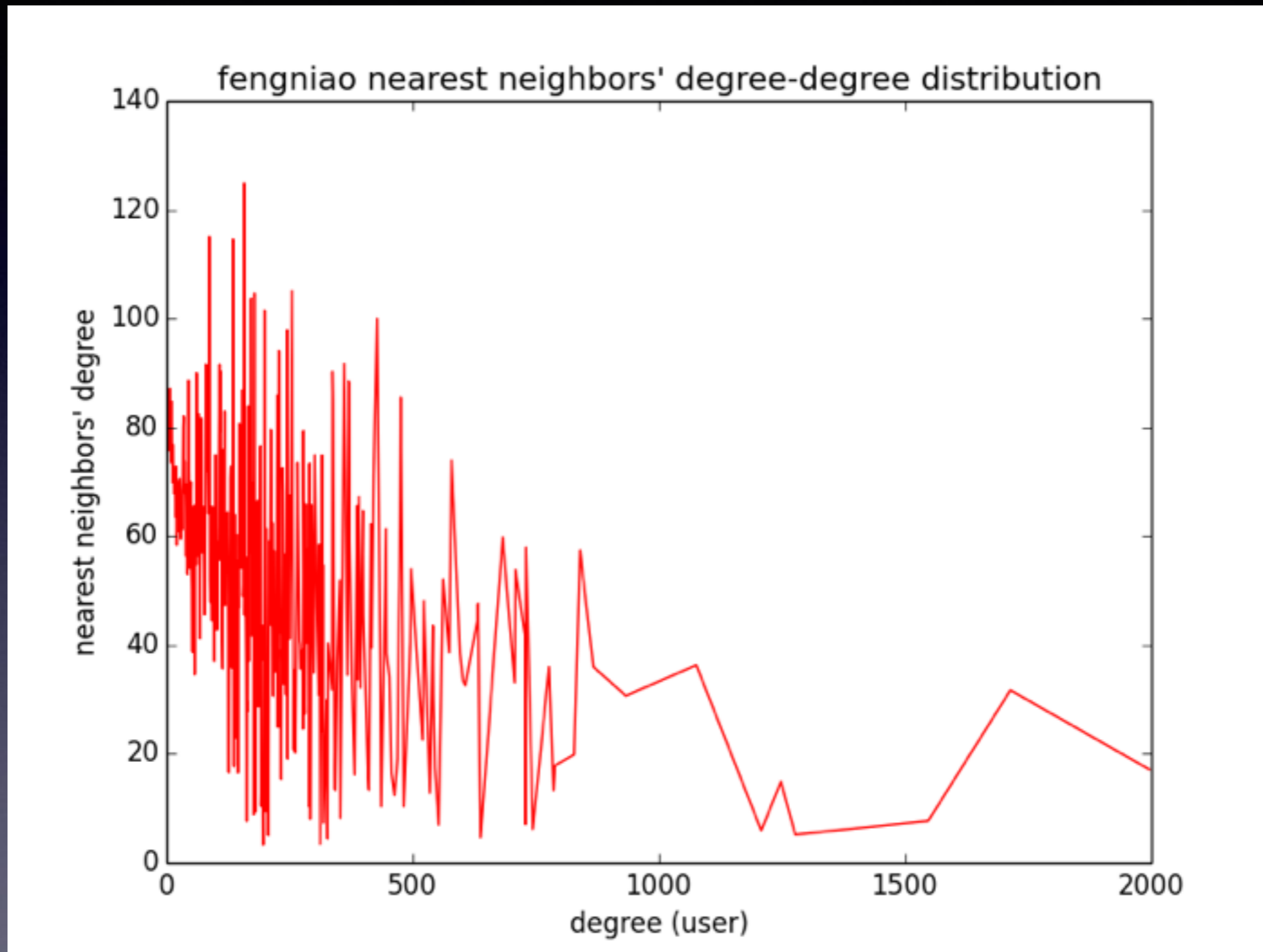
Fresh (small degree) users tend to select popular (large degree) objects.[1]

Nearest Neighbors' Degree (Object)

Unpopular (small degree) objects are more likely be selected by active (large degree) users.[1]

How users' interests evolve over time?

How to measure users' interests?
we measure it via **average similarity among objects** collected by the user which is exactly the collaborative similarity we mentioned above.

How to measure users' basic interests?
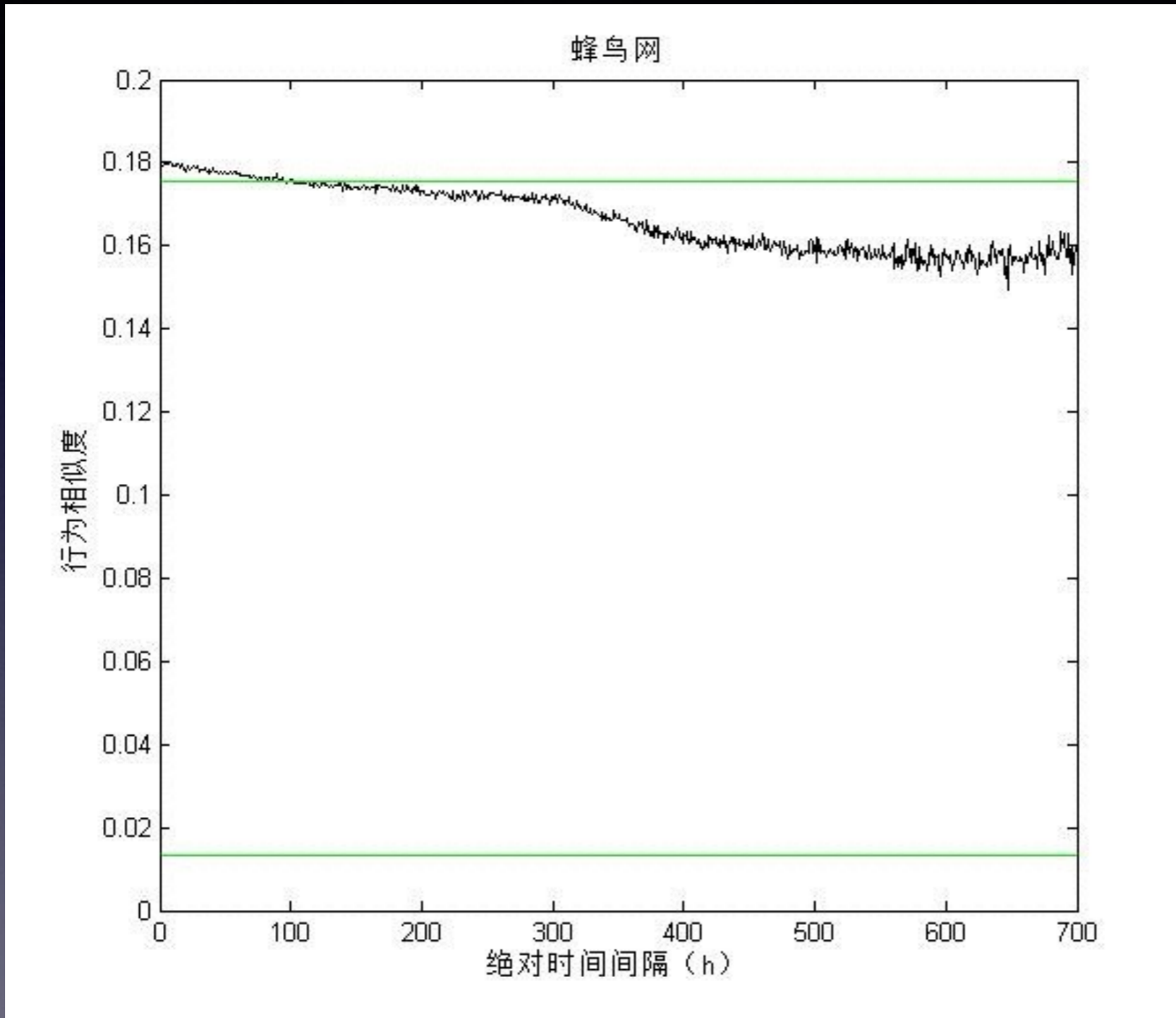Via **average similarity among all the objects** collected by the user.

How to measure users' temporal interests?
Via **average similarity among the objects** collected in a **specific time difference** by the user.

We calculate **the average similarity among all the objects** as a baseline.
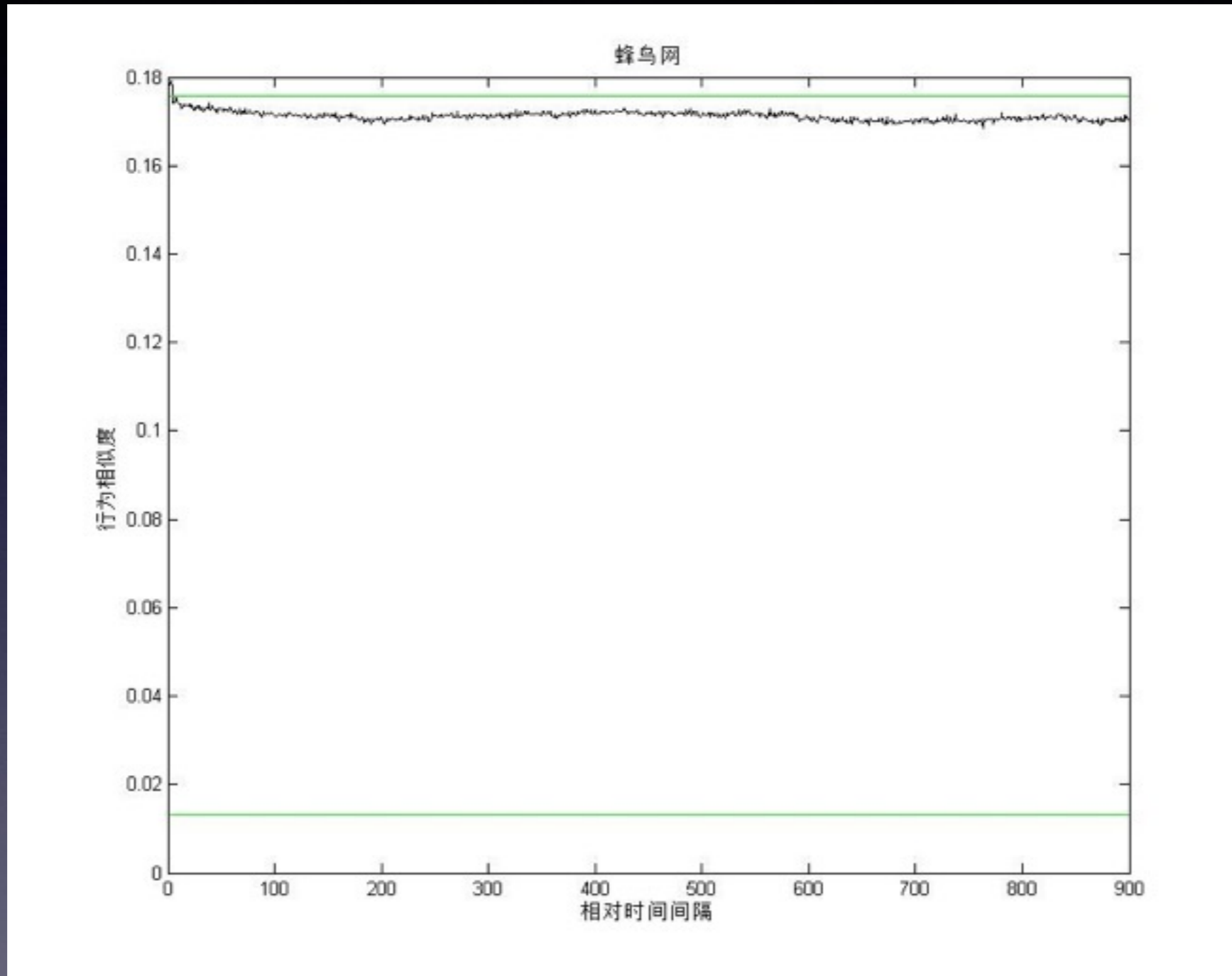
Interests - Absolute Time

Interests

Absolute time (physical time)

Interests - Relative Time

Interests

Relative time (ordinal time)

Interests - Summary

The upper green line indicates users' basic interests.

The lower green line indicates average similarity among all the objects.

The middle black curve indicates how users' interests evolve over time.

References

[1]   Shang, Ming-Sheng, et al. "Empirical analysis of web-based user-object bipartite networks." *EPL (Europhysics Letters)* 90.4 (2010): 48006.