

# AIDE: An Accurate Internet Distance Estimation System



Chengbo Dong, Guodong Wang, Beixing Deng and Xing Li  
Tsinghua National Laboratory for Information Science and Technology  
Department of Electronic Engineering, Tsinghua University, Beijing 100084, P.R. China

## 1. Motivation

- IDES is the first matrix factorization based network coordinate (NC) system.
- However it suffers from a significant drawback: high distance estimation error in short links.
- This significant drawback degrades practical performance in many Internet applications, such as network overlay multicast and server selection.

## 2. Aim

We aim to propose an accurate Internet distance estimation system, AIDE to increase the relative errors of short links. Therefore, many Internet practical applications will benefit from it.

## 3. AIDE Design

AIDE is based on singular value decomposition.

### Principle of AIDE system

The first  $L$  joining hosts are called landmarks. AIDE first uses SVD algorithm to factorize the initial landmark distance matrix to get the singular value, which is approximately suitable for the whole network distance matrix  $D$ .

### Work Flow of AIDE

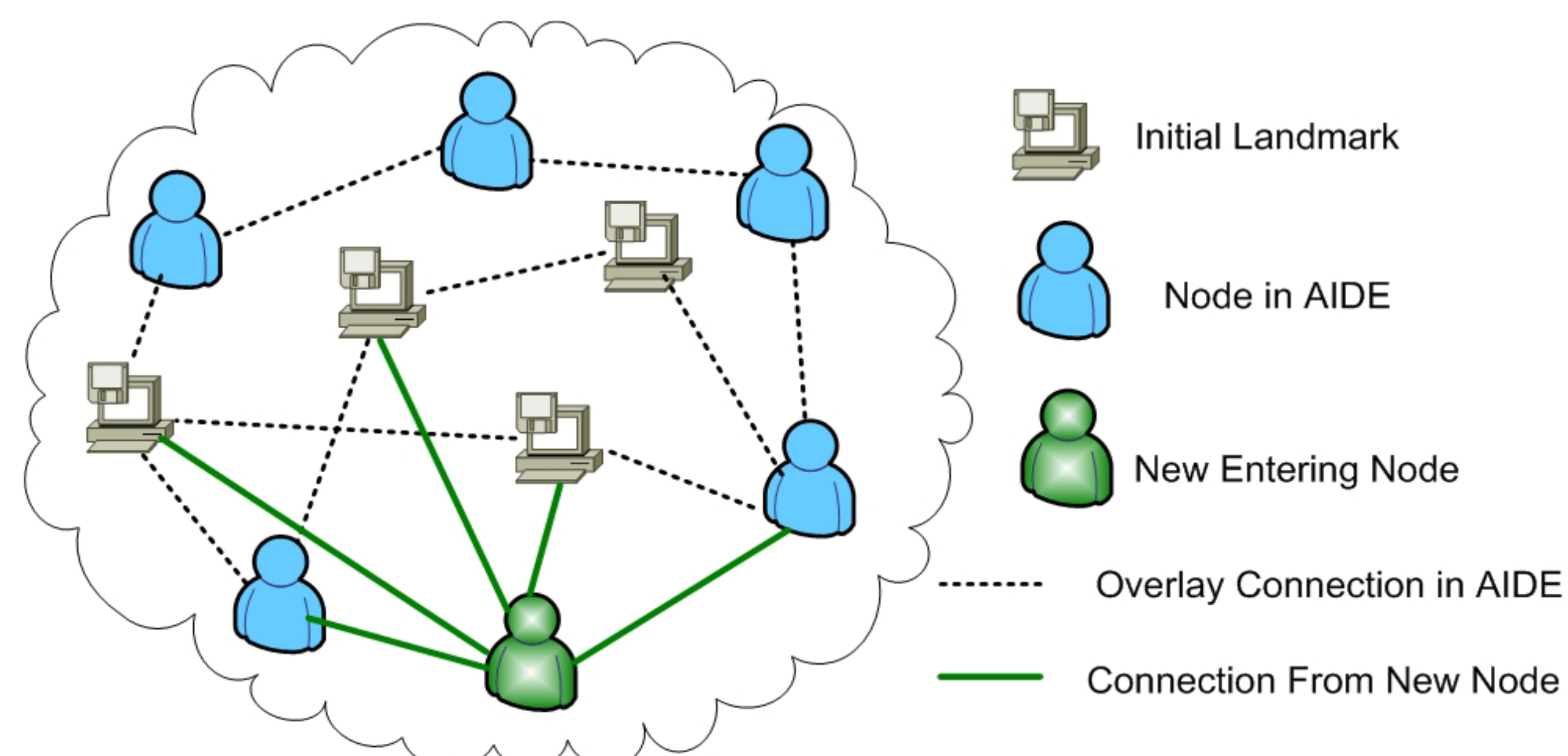


Fig. 1. AIDE system design

When host  $H_i$  is computing its vectors in AIDE system, it can select the initial landmarks or any hosts that have joined the system as reference nodes. For example, it selects  $k$  initial landmarks and  $L - k$  closest neighbor nodes in the system as its  $L$  reference nodes  $R_{i1}, \dots, R_{iL}$ .

### How to reduce short distance estimation errors?

Selecting the closest hosts as landmarks can reduce the short link estimation errors. However, choosing too many new joined hosts as landmarks may cause the decreasing of the relativity between vectors of the initial landmarks and the singular values.

### Tradeoff between reducing short link estimation errors and keeping the relativity between vectors of the initial landmarks and the singular values.

In our experiments, we assign 20 reference nodes for each joining hosts, with 15 initial landmarks and 5 closest neighbors.

## 4. Metric

### Relative Error (RE)

$$RE = \frac{|MeasuredDistance - PredictedDistance|}{\min(MeasuredDistance, PredictedDistance)}$$

This metric describes the estimate accuracy of a NC system.

### Nearest Neighbor Loss (NNL)

Nearest neighbor loss (NNL) is defined as the difference between the delay to the node predicted to be the nearest neighbor and the delay to the true one. This metric reflects the accuracy for one host node in a network coordinate system to find the closest neighbor host.

## 8. Acknowledgements

This work is supported by the National Science Foundation of China (No. 60850003, No.60473087, No.60703052) and the National Basic Research Program of China (No.2007CB310806).

## 5. Evaluation

Experiment Setup:

- King data set with 1740 nodes
- Comparing AIDE with IDES
- Assign 20 reference nodes in both simulations.

### A. Relation between relative error and range of distance

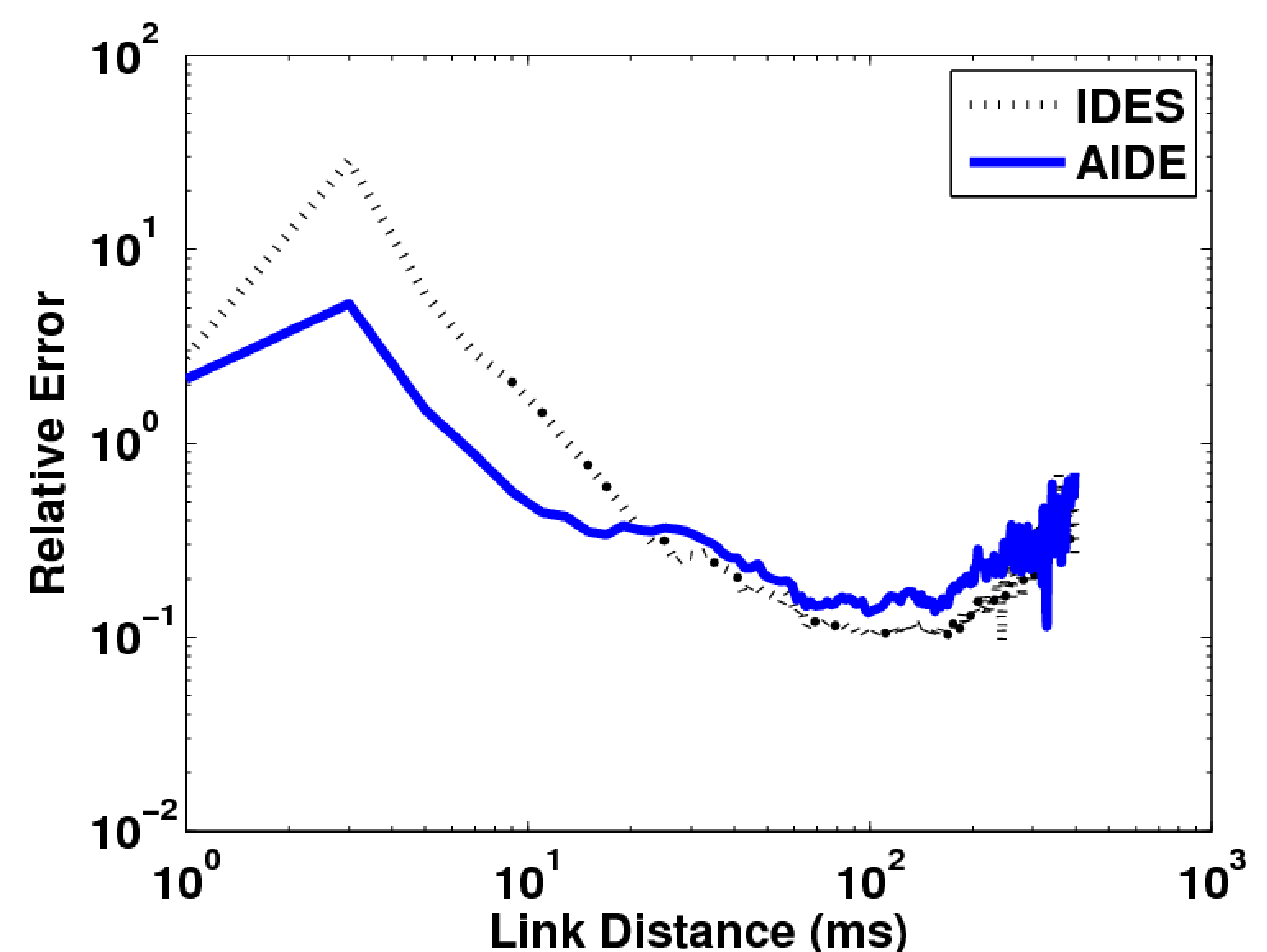


Fig. 2. Relation Between RE and Link Distance

Figure 2 shows that IDES system suffers from high relative error of short link distance, while AIDE has better performance in short link estimation with more accuracy.

### B. Nearest Neighbor Loss

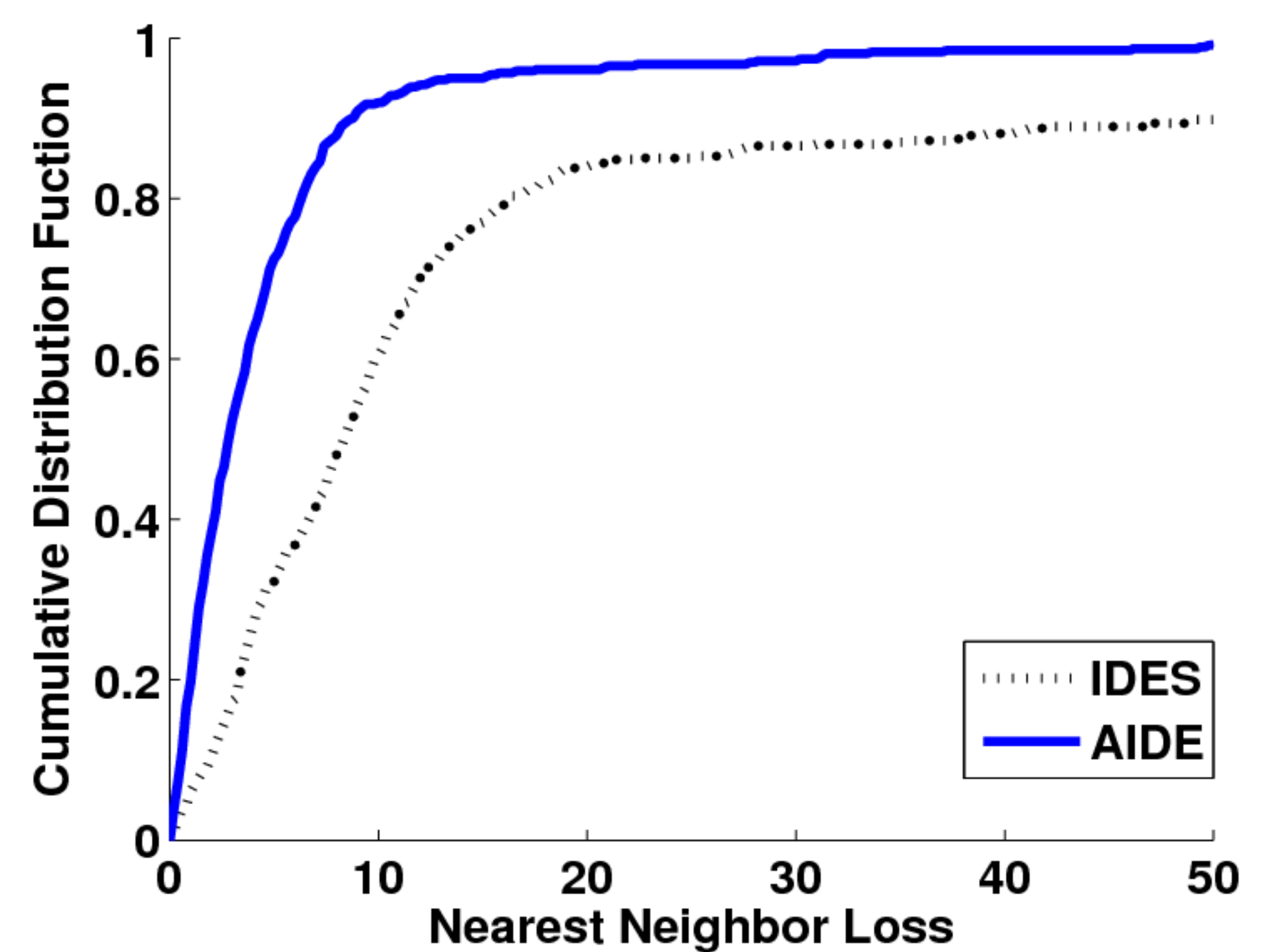


Fig. 3. Cumulative Distribution Function of NNL

Figure 3 shows the NNL of AIDE and IDES. It is clearly obvious that AIDE has much better performance than IDES in this metric which means AIDE predicts the distance between the closest hosts with far more accuracy. We believe this advantage will benefit Internet applications.

## 6. Conclusion

- AIDE is a novel accurate and decentralized NC system based on matrix factorization.
- AIDE significantly outperforms IDES in estimation of short link distance, which will bring great benefits in many applications.
- Since AIDE does not rely much on infrastructure nodes, AIDE is a scalable and robust NC system with accurate estimation in large-scale network distance.

## 7. Future Work

We plan to design a practical NC system using AIDE algorithm.