

ABSTRACT

DISTRIBUTED GRAPH EMBEDDING METHOD AND APPARATUS, DEVICE, AND SYSTEM

The embodiments of the present specification provide a distributed graph embedding method and
5 apparatus, a device, and a system. A graph data structure is distributed in at least two computing
nodes. The method includes: when a preset sampling algorithm is used to perform path sampling on
the graph data structure, adding an identification and a current embedding vector of a current sampled
vertex to a predefined message data structure; when a data length of the message data structure
reaches a window length defined by a word vector algorithm, using the word vector algorithm to
10 calculate an update gradient of an embedding vector of each sampled vertex in an updated message
data structure; based on identifications in the updated message data structure, sending each update
gradient to a corresponding vertex respectively, and using the update gradients to update the current
embedding vector of the sampled vertex belonging to the current computing node; and after a preset
convergence condition is met, constituting a graph embedding matrix by final embedding vectors of
15 all vertices in the graph data structure.

[Fig. 2A]