

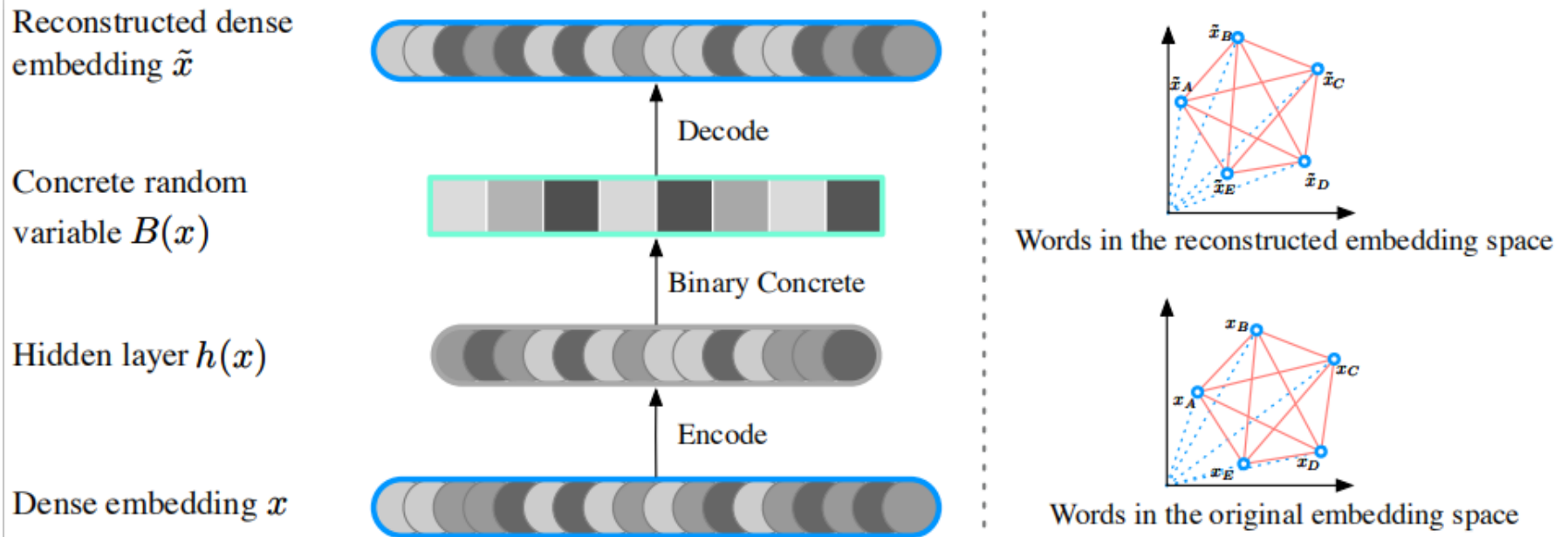
# Relation Reconstructive Binarization of Word Embeddings

**Feiyang PAN, Shuokai LI, Xiang AO, Qing HE**

Frontiers of Computer Science, DOI: [10.1007/s11704-021-0108-3](https://doi.org/10.1007/s11704-021-0108-3)

# Problems & Ideas

- Problems of previous word-embedding compression methods
  - They only aim to preserve most information of every individual word, and often fail to retain the relation between words.
- Ideas: Adaptive virtual parallel TCP
  - We proposed to transform word embeddings into binary codes that can preserve the relation between words.



# Main Contributions

- **We propose Relation Reconstructive Binarization (R2B), which trains an auto-encoder towards minimizing the reconstruction error of word-by-word relations.**
- **To efficiently learn the word relations, we propose a training method which has a linear time and space complexity with respect to the vocabulary size.**
- **Extensive experiments on both word semantic tasks and downstream machine learning tasks (text classification, sentence matching and machine translation) demonstrate the superiority of R2B.**