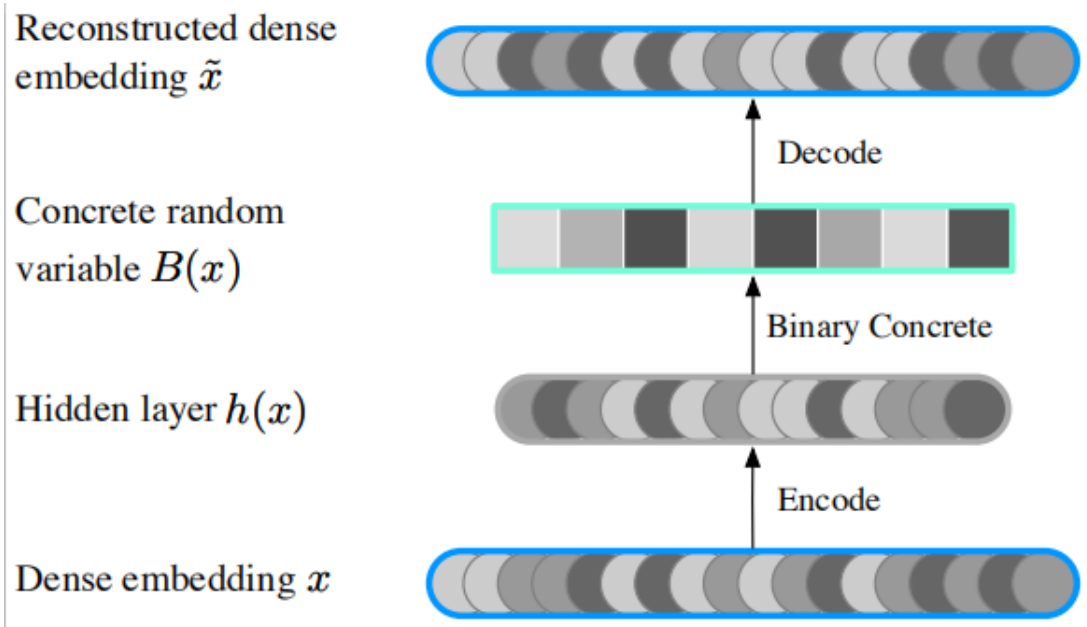
# Relation Reconstructive Binarization of Word Embeddings

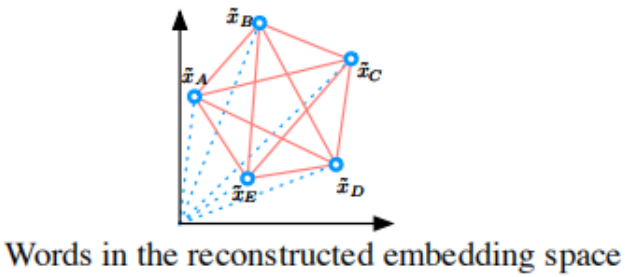
**Feiyang PAN, Shuokai LI, Xiang AO, Qing HE**

Manuscript ID: FCS-20108

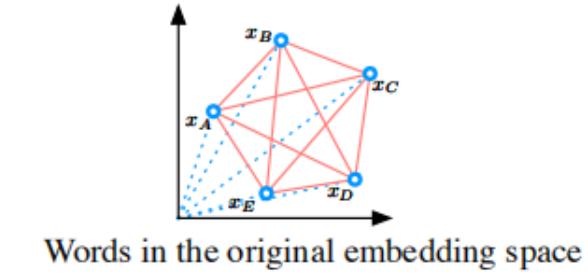
# Problems & ideas

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



Words in the reconstructed embedding space



Words in the original embedding space

# 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.**