AliGraph: An Extremely Large Scale Graph Representation Learning in Practice

Xiaoyong Liu
Hongxia Yang
Wei Lin

Alibaba Group
Part 1: About Alibaba
Our Mission

让天下没有难做的生意
To make it easy to do business anywhere
Our Vision

MEET
Enabling millions of commercial and social interactions

WORK
Empowering our customers with data and infrastructure to manage their business

LIVE
To become central to the everyday lives of our customers

@ Alibaba
Our Marketplaces are Networked

Cross Promotion Across Our China Retail Marketplaces

Strong Synergies Between Our Retail & Wholesale Marketplaces
Part 2 : Why GE and Two works
Why Graph and Graph Embedding?

- Graph computing models are very popular in big data companies, especially IT companies, as they are the most straightforward solutions to many practical problems.

- Traditional recommendation and CTR/CVR prediction problems can be equivalently modeled with the attributed user-item bipartite graphs.

- Objective functions are more general: global optimization vs conditional independent optimization.

- Introducing high-level proximity samples and their modeling brings both risks (e.g., noise) and benefits (e.g., more generalization and exploration, predictive graph changes).

- Pure deep learning is mature, graph embedding that combines both deep learning and graph computing integrating end-to-end learning with inductive reasoning, which is expected to solve the relational reasoning that deep learning cannot perform[1].

Representative Work: Graph Embedding in Fraud Detection

Challenges of mobile fraud detection:
- Billions of mobile logging records per day
- Device ids are missing
- Devices are abnormal:
  - Dual sim cards dual standby: one device has two imei, imsi[2] with random switching
  - Replace the sim card: one device corresponds to multiple imsi
  - System Restore: One device corresponds to multiple device ids
  - Simulator: One device corresponds to a large number of device ids
  - Cottage machine: a large number of devices share one same id
- Devices contain rich attributes

We propose a distributed deep learning-based representation learning model, ANRL, which maps nodes in the graph to representation vectors in low-dimensional space to facilitate subsequent processing of related tasks
- Combining both network structure and node attribute information to better maintain network characteristics
- Neural networks can mine deeper correlations between the two

[1] ANRL: Attributed Network Representation Learning via Deep Neural Networks, IJCAI 2018
[2] imei: international mobile equipment identity; imsi: international mobile subscriber identity
• **Objective Functions:**

\[ \mathcal{L} = \mathcal{L}_{reg} + \alpha \mathcal{L}_{ae} + \beta \mathcal{L}_{reg} \]
\[ = - \sum_{i=1}^{n} \sum_{c \in C} \sum_{-b \leq j \leq b, j \neq 0} \log \frac{\exp(u_i^T y_i^{(K)})}{\sum_{v=1}^{n} \exp(u_v^T y_i^{(K)})} \]
\[ + \alpha \sum_{i=1}^{n} \| \hat{x}_i - T(v_i) \|^2_2 + \frac{\beta}{K} \sum_{k=1}^{K} (\| W^{(k)} \|^2_F + \| \hat{W}^{(k)} \|^2_F) \]

• Existing methods generally incorporate attribute information into representation learning by matrix decomposition or by adding auxiliary nodes. This shallow level model can’t capture the deep inner relationship between the two.

• Compared with the existing state-of-the-art method, our model ANRL has a certain degree of improvement in tasks such as ink prediction and node classification by using the deep learning model.
Representative Work: Graph Embedding in Entity Recognition[1]

Challenges of mobile entity recognition:
• Data contains large-scale sparse network structure information and rich attribute information
• Both information can be used to describe the relationship among the nodes, however distance measures of the two are probably not monotonically increasing or decreasing of the same nodes
• Consider a non-positive relationship, we can capture more information to help improve the learning performance

We propose to investigate attributed network embedding through taking uncorrelation between topology structure and attribute into account and make the following contributions.
• We recognize the uncorrelation phenomenon between topology structure and attribute which affects the actual proximity among nodes in the embedding space.
• We propose a Personalized node Relation Ranking Embedding (PRRE) model such that the resulting network embedding can capture the uncorrelation as well as preserve good proximity among nodes.
• Sampling with Mini-batch Gradient Descent for efficient iteration and update.

Objective function and algorithm framework

\begin{align*}
\mathcal{J}(H, \theta_T, \theta_A) & = \prod_{i \in V} \left( \prod_{p \in P} \prod_{a \in A} P(p \geq 1 | a; \theta_A, \theta_T, H) \cdot \prod_{a \in A} \prod_{n \in N} P(a \geq 1 | n; \theta_A, \theta_T, H) \right).
\end{align*}

\begin{align*}
\mathcal{J}(H, \theta_T, \theta_A) & = \ln \mathcal{J}(H, \theta_T, \theta_A) - \lambda_h \sum_{i \in V} \| h_i \|^2 \\
& = \sum_{i \in V} \left( \sum_{p \in P} \sum_{a \in A} \frac{1}{1 + g(\theta_T, \theta_A)} \ln \sigma_{ip} - \sigma_{ia} + 1 \right) + \\
& \sum_{a \in A} \sum_{n \in N} \frac{1}{1 + g(\theta_T, \theta_A)} \ln \left( \frac{\sigma_{ia} - \sigma_{in} + 1}{2} \right) - \lambda_h \sum_{i \in V} \| h_i \|.
\end{align*}

**Algorithm 1** PRRE for attributed networks

**Input:** $G = \{ V, E, A \}, d$

**Output:** $H \in \mathbb{R}^{n \times d}$

1. Compute similarity matrix $S_A \in \mathbb{R}^{n \times n}$ and $S_T \in \mathbb{R}^{n \times n}$ using selected similarity measure
2. Initialize $\theta_A, \theta_T$ and $H \sim U(0, 1)$
3. while $t < \text{max\_iter}$ and $\Delta \mathcal{J} < \epsilon$ do
4. Sample batch $B$ with size $s$,
5. for $v_i \in V$ do
6. Compute Positive/Ambiguous/Negative pairs using current thresholds $\theta_T, \theta_A$
7. end for
8. Compute the gradients for $h_i, h_p, h_a, h_n$ by Equation [13], [14], [15], [16].
9. Update $H$ by Equation [19]
10. Compute the gradients for $\theta_T, \theta_A$ by Equation [17], [18].
11. Update $\theta_T, \theta_A$ by Equation [19]
12. end while
Other Graph Embedding Representative Works

5. Mobile access record resolution on large-scale identifier-linkage graphs. KDD, 2018.
8. Heterogeneous Embedding Propagation for Large-scale E-Commerce User Alignment, 2018 IEEE International Conference on Data Mining (ICDM), 2018
12. Will Triadic Closure Strengthen Ties in Social Networks?, ACM Transactions on Knowledge Discovery from Data (TKDD), 2017.
Part 3 : PAI - AliGraph Algorithm System
PAI: Alibaba Machine Learning Platform of AI
PAI Architecture

AI Solutions (Recommendation, OCR, NLP, Robot, Vision)

- Smart Labeling
  - Multi Scenarios
  - Standard Templates
  - Data Management

- PAI-EAS, Elastic Algorithm Service
  - Cloud Native
  - High Performance
  - Elastic
  - Build-in Accelerator

- PAI-Studio
  - 100+ Algorithm Components
  - Distributed
  - AutoML

- PAI-DSW
  - Various Frameworks, TF, PyTorch...
  - JupyterLab, WebIDE, Terminal

- PAI-DLC
  - Cloud Native
  - Easy to use
  - High performance

- PAI AutoLearning
  - One-Stop System
  - Transfer Learning
  - Model Zoo
  - Labeling System

ML Framework (Alink / MPI / PS / Graph / TensorFlow/PyTorch/Caffe...)

Engines (MaxCompute / EMR / Flink)

ACK (Alibaba Kubernetes)

Infrastructure (CPU, GPU, FPGA, NPU)

https://www.alibabacloud.com/product/machine-learning
**PAI-AliGraph Algorithm Framework**

**Algorithm 1: GNN Framework**

**Input:** network $G$, embedding dimension $d \in \mathbb{N}$, a vertex feature $x_v$ for each vertex $v \in \mathcal{V}$ and the maximum hops of neighbors $k_{\text{max}} \in \mathbb{N}$.

**Output:** embedding result $h_v$ of each vertex $v \in \mathcal{V}$

1. $h_v^{(0)} \leftarrow x_v$
2. for $k \leftarrow 1$ to $k_{\text{max}}$
   3. for each vertex $v \in \mathcal{V}$ do
      4. $S_v \leftarrow \text{SAMPLE}(N_b(v))$
      5. $h_v' \leftarrow \text{AGGREGATE}(h_u^{(k-1)}, \forall u \in S)$
      6. $h_v^{(k)} \leftarrow \text{COMBINE}(h_v^{(k-1)}, h_v')$
   7. normalize all embedding vectors $h_v^{(k)}$ for all $v \in \mathcal{V}$
8. $h_v \leftarrow h_v^{(k_{\text{max}})}$ for all $v \in \mathcal{V}$ return $h_v$ as the embedding result for all $v \in \mathcal{V}$

- Business needs to carry out more model innovations based on the AliGraph Algorithm Framework

- Need more flexible programming framework to support algorithm innovation
  - PAI Tensorflow
AliGraph-System Overview

- Graph storage + sampling + operator need to be tightly combined for deep learning together

- Aligraph Algorithm Framework based on PAI-Tensorflow extension
  - Use TF's flexible composition method
  - Use TF's good design structure for effective expansion
  - Most of the existing algorithm innovations are based on TF
AliGraph-Graph Building
• Various sampling modes require flexible support
• Provide traverse, neighborhood, negative three ways
• Overlap sampling and calculation

```python
def sampling(s1, s2, s3, batch_size):
    vertex = s1.sample(edge_type, batch_size)
    context = s2.sample(edge_type, vertex, hop_nums)
    neg = s3.sample(edge_type, vertex, neg_num)
    return vertex, context, neg
```
Graph Storage (Memory Representation)

- Tens of billions of edge relations
- The adjacency relationship is the basis of the GNN algorithm, and locality needs to be fully utilized
- Tradeoff between data fragmentation and communication
  - Effective caching strategy: importance metric based on vertices
- Easy to traverse and sample by batch
Cache/Memory Strategy

\textbf{Input:} graph $G$, partition number $p$, cache depth $h$, threshold $\tau_1, \tau_2, \ldots, \tau_h$

\textbf{Output:} $p$ subgraphs

1. Initialize $p$ graph servers
2. For each edge $e = (u, v) \in E$ do
   3. $j = \text{ASSIGN}(u)$
   4. Send edge $e$ to the $j$-th partition
3. For each vertex $v \in V$ do
   4. For $k \leftarrow 1 \text{ to } h$ do
      5. Compute $D^{(k)}(v)$ and $D^{(k)}_o(v)$
      6. If $\frac{D^{(k)}(v)}{D^{(k)}_o(v)} \geq \tau_k$ then
         7. Cache the 1 to $k$-hop out-neighbors of $v$ on each partition where $v$ exists

---

Cache acceleration: 50% faster than random method, 60% faster than LRU method
Operator Optimization

- Reduce double calculation by caching all levels of $h$
- Co-locate model parameters, all levels of $h$ and the graph itself, greatly reduce network communication and delay
- Algorithm and System co-design: need to cooperate with training strategy and upper-level algorithm
PrepareData
BuildGraph
GSL SampleStream
FeatureVectorization
GNN Module
DefineLossFunction
TrainingIteration
ExportModule/Vector

```python
# Define graph object
g = gl.Graph()

# Add data source
g.node(1_path, '1', decoder=gl.Decoder(attr_types=['float'] * 4, attr_dims=[10] * 4, labeled=True))
.edge(12l_path, ('1', '1', '1-1'), decoder=gl.Decoder())
.init()

# Construct GSL Query
query = g.V('i').batch(10).alias('1')
    .outV('1-1').sample(5).by('topk').alias('hop1')
    .outV('i-1').sample(5).by('random').alias('hop2')
    .values()
df = tfm.DataFlow(query)

# Construct Module
dims = np.array([4, 16, 8])
model = tfm.HomoEgoGraphSAGE(dims, bn_fn=None, active_fn=tf.nn.relu, dropout=0.1)

# Module computing, Generate embedding
embeddings = model.forward(df.get_ego_graph('1'))

# Construct Node Classifier
nc = tfm.NodeClassifier(dims=[8, 4], class_num=2)
logits, loss = nc.forward(embeddings, ex.nodes.labels)

# Train
trainer = tfm.Trainer()
trainer.minimize(loss)

def trace(ret):
    print('loss = %f' % ret[1])
trainer.step_to_epochs(10, [logits, loss], trace)
g.close()
```
AliGraph-Algorithm Deployment and Share

- PAI provides algorithm warehouse for algorithm package, management and release
Part 4: Algorithm Warehouse

- Representative and negative sampling
- Borrow idea from importance sampling
- Extend node-wise to batch-wise sampling
- Type-dependent & -fusion sampling with self-normalization
- Computational complexity drops from $O(|E|+|V|)$ to $O(|V|)$
Table 4: Micro/Macro-F1 scores for multi-class classification on Aminer. Excluding HEP, the best method is bolded and the second best is underlined. Percentages in parenthesis indicate the performance level, treating HEP-nil as 0% and HEP as 100%.

<table>
<thead>
<tr>
<th>Sampling size</th>
<th>Micro-F1</th>
<th>Macro-F1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>512</td>
<td>1024</td>
</tr>
<tr>
<td>HEP-Nil</td>
<td>~3%</td>
<td>0.2084 (20%)</td>
</tr>
<tr>
<td>HEP</td>
<td>0.9566 (95%)</td>
<td>0.9551 (95%)</td>
</tr>
<tr>
<td>AS-GCN</td>
<td>0.2866 (10%)</td>
<td>0.2901 (11%)</td>
</tr>
<tr>
<td>Unif-TD</td>
<td>0.3089 (13%)</td>
<td>0.3934 (25%)</td>
</tr>
<tr>
<td>Unif-TF</td>
<td>0.2877 (11%)</td>
<td>0.4010 (26%)</td>
</tr>
<tr>
<td>Unif-TD-SN</td>
<td>0.2283 (3%)</td>
<td>0.2955 (12%)</td>
</tr>
<tr>
<td>Unif-TF-SN</td>
<td>0.1631 (-6%)</td>
<td>0.2951 (12%)</td>
</tr>
<tr>
<td>VarR-TD</td>
<td>0.3423 (18%)</td>
<td>0.4911 (38%)</td>
</tr>
<tr>
<td>VarR-TF</td>
<td><strong>0.6200 (55%)</strong></td>
<td>0.7472 (72%)</td>
</tr>
<tr>
<td>VarR-TD-SN</td>
<td>0.2353 (4%)</td>
<td>0.3245 (16%)</td>
</tr>
<tr>
<td>VarR-TF-SN</td>
<td><strong>0.5960 (52%)</strong></td>
<td><strong>0.8124 (81%)</strong></td>
</tr>
</tbody>
</table>

Table 5: F1 and AUC scores for binary purchase prediction on Alibaba. Excluding HEP, the best method is bolded and the second best is underlined. Percentages in parenthesis indicate the performance level, treating HEP-nil as 0% and HEP as 100%.

<table>
<thead>
<tr>
<th>Sampling size</th>
<th>F1-score</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>512</td>
<td>1024</td>
</tr>
<tr>
<td>HEP-Nil</td>
<td>~2.5%</td>
<td>0.3994 (intuitive lower bound)</td>
</tr>
<tr>
<td>HEP</td>
<td>0.5793 (intuitive upper bound)</td>
<td>0.7777 (intuitive upper bound)</td>
</tr>
<tr>
<td>AS-GCN</td>
<td>(unable to complete due to memory constraint)</td>
<td></td>
</tr>
<tr>
<td>Unif-TD</td>
<td>0.3883 (-6%)</td>
<td>0.4353 (20%)</td>
</tr>
<tr>
<td>Unif-TF</td>
<td>0.4051 (3%)</td>
<td>0.4281 (16%)</td>
</tr>
<tr>
<td>Unif-TD-SN</td>
<td>0.3884 (-6%)</td>
<td>0.4322 (18%)</td>
</tr>
<tr>
<td>Unif-TF-SN</td>
<td>0.3928 (-4%)</td>
<td>0.4219 (13%)</td>
</tr>
<tr>
<td>VarR-TD</td>
<td>0.4476 (27%)</td>
<td>0.4518 (29%)</td>
</tr>
<tr>
<td>VarR-TF</td>
<td><strong>0.4774 (43%)</strong></td>
<td><strong>0.4863 (48%)</strong></td>
</tr>
<tr>
<td>VarR-TD-SN</td>
<td>0.4293 (17%)</td>
<td>0.4497 (28%)</td>
</tr>
<tr>
<td>VarR-TF-SN</td>
<td>0.4766 (43%)</td>
<td><strong>0.4896 (50%)</strong></td>
</tr>
</tbody>
</table>
Figure 1: The left illustrates an example of an attributed multiplex heterogeneous network. Users in the left part of the figure are associated with attributes including gender, age, and location. Similarly, items in the left part of the figure include attributes such as price and brand. The edge types between users and items are from four interactions, including click, add-to-preference, add-to-cart and conversion. The three subfigures in the middle represent different ways of setting up the graphs, including HON, MHEN, and AMHEN from the bottom to the top. The right part shows the performance improvement of the proposed models over DeepWalk on the A+ dataset. As can be seen, GATNE-I achieves a +28.23% performance lift compared to DeepWalk.
• Transudative Model:

For GATNE-T, the overall embedding for node \( v_i \) on edge type \( r \) is:

\[
v_{i,r} = b_i + \alpha_r M_r^T U_i a_{i,r} = b_i + \alpha_r M_r^T \sum_{p=1}^{m} \lambda_p u_{i,p},
\]

where \( \lambda_p \) denotes the \( p \)-th element of \( a_{i,r} \) and is computed as:

\[
\lambda_p = \frac{\exp(w_r^T \tanh(W_r u_{i,p}))}{\sum_t \exp(w_r^T \tanh(W_r u_{i,t}))}.
\]

• Inductive Model:

\[
v_{i,r} = h_z(x_i) + \alpha_r M_r^T U_i a_{i,r} + \beta_r D_z^T x_i,
\]
Figure 3: (a) The convergence curve for GATNE-T and GATNE-I models on A+ dataset. The inductive model converges faster and achieves better performance than the transductive model. (b) The training time decreases as the number of workers increases. GATNE-I takes less training time to converge compared with GATNE-T.

\[ p(o|s(o), \mathcal{P}, \theta) = \prod_{v_j \in N(v_i)} p(v_j|v_i, s(o)), \]

and each product factor is calculated as

\[ p(v_j|v_i, s(o)) = \frac{\exp(<H_j^k, U_i^k>)}{\sum_{v, k} \exp(<H_i^k, U_i^k>)}, \]
Figure 2: Framework of Bi-HGNN. Left side outputs user-community and individualized embeddings. Since user and item raw features are generally sparse, wide and deep is used here for feature encoding. Based on the user’s historical behaviors, $n$ items are selected as his/her neighbors to generate full user embedding with GraphSAGE. User-community embedding is derived through weighted average over users that belong to the specific community, where weight is determined by the distance between the user and base community embedding. Full embedding is decomposed into user-community and individualized embeddings, which is illustrated in the upper left of the figure within the dashed square. Right part represents the item embedding generation. With both the user and item embeddings, the classifier undergoes several fully connected neural networks and outputs a probability to determine whether to recommend the item to the user.
Figure 2: An illustration of the proposed framework *BurstGraph*. At time step $t$, the framework generates vanilla evolution and bursty evolution based on network structure $G_t$. Part a is an original VAE for vanilla evolution, where random variable $z_t$ follows a Gaussian distribution. Part b is an extended VAE for bursty evolution, where random variable $s_t$ follows a spike-and-slab distribution because of the sparsity of bursty links. The encoder for these two random variables $z_t$ and $s_t$ shares the same GraphSAGE to utilize the information from vertices and their neighbors.
Bayesian GNN

- BEM is proposed to bridge KG and BG seamlessly, with the consideration of the behavior-specific bias. This framework provides a new perspective of making a reasoning mechanism (cognitive graph).
- As a method, BEM is generic and flexible in that it can use any KG embeddings to correct any BG embeddings. On the contrary, it is potentially able to help KG embeddings acquire novel knowledge from the BG embeddings that does not exist in the knowledge graph.

![Reconstructed vertex graphs](image)

Figure 4: Reconstructed vertex graphs. The blue ones are original graphs and red ones are corrected graphs.
## Supported Algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Heterogeneous</th>
<th>Attributed</th>
<th>Dynamic</th>
<th>Large-Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Node</td>
<td>Edge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>AHEP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>GATNE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Mixture GNN</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Hierarchical GNN</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Bayesian GNN</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Evolving GNN</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Part 5: Current Focus in Practice
Search-Recommendation Platform

**Business Solution**
- Search/Recommendation/Ads
- Business Intelligent
- Fraud Detection/Risk Control
- ID Mapping and Profile
- Crowd Marketing

**Service**
- Index/Rank
- Training/Inference
- Evolutonal/Online updating
- AutoML
- Transfer Learning/Reinforcement Learning

**AI Algorithm**
- Natural Language Understanding
- Image Processing
- Video Processing
- Machine Learning/Deep Learning
- Cognitive Graph

**Data**
- Global User Profile
- Global Seller Profile
- Online Offline Info Integration
- Ecommerce Cognitive Graph
- Business Operation Knowledge

**System**
- Distributed Computing
- Stream Computing
- Graph Computing
- Online Learning and Inference
- Heterogeneous Computing
Cloud Theme

from Single Product Recommendation to Group Meet Place, cultivate metal upgrade

Meeting Place for Personalized Groups

- SPR
- Reasoning
- Tips
- Knowledge Card
- Video recommendation

Ad Embedding On Scene

Algorithm

- Natural Language Understanding
- Image Processing
- Video Processing
- Machine Learning/Deep Learning
- Graph Embedding
Cloud Theme The Flow

**Stage 1**

- **Product-Side**
  - GNN Model Generation Embedding
  - Tree Based Clustering Learning user interest category
  - Mining Corresponding category keywords/phrase
  - Generate Personalized themes and images

- **User-Side**
  - Embedding Recall Related Users
  - Group personalization User Profile mining

**Stage 2**

- Personalized cloud theme Knowledge card and strategy generation
1. GCN Model generate embedding
2. Hierarchical clustering, each layer is divided
3. As the number of layers deepens, the granularity of the division becomes finer, and the semantics gradually emerge
Cloud Theme Title Auto-Generation

Now:
1. thousands of online cloud themes, mostly manual verification, many unmaintained
2. The main and subtitles of the cloud theme are too plain to attract users to click

Future:
1. Automatically generate cloud topics through algorithms, produce them in batches by algorithms, and automatically iteratively optimize.
2. Learn to produce scene-oriented, attractive main and subtitles.
   ex:
   1st cloud theme
   Main Title: The coolest "digital" baby
   Subtitle: Immersive game experience

   2nd cloud theme
   Main Title: Baby winter Heater
   Subtitle: Provide warm sleep for your baby
**Cloud Theme Title Auto-Generation**

**Method 1: retrieval model**

- **Civic 10th gen modification**
- **Civic Modification**
- **Angkerra Modification**
- **Fit modification**

---

**Appearance of 10th-gen Civic is modified to make it even more exciting.**
- The new Civic changes, **turn your car into a wild horse.**
- **Nothing is born perfect**, a Civic modified car fits you.
- The new Civic is modified to **play a high-powered style.**
- The 205-horsepower grocery shopping car, the Civic is so handsome.
- Upgrade the Civic to **create a savage calf.**

---

**Turn your car into a wild horse**
- Make the car more exciting
- **Nothing is born perfect**
- create a savage calf

---

**IDF Values**
- 0.4532
- 0.3312
- 0.2545
- 0.1124

---

**Sorted Result Set**
## Cloud Theme Title Auto-Generation

<table>
<thead>
<tr>
<th>序列</th>
<th>主题</th>
<th>副标题</th>
<th>类目</th>
</tr>
</thead>
<tbody>
<tr>
<td>1019</td>
<td>你的装嫩神器</td>
<td>完美发挥显示器的性能</td>
<td>显示器 键盘 散热器/风扇 Hifi音箱 机箱 DIY兼容机 鼠标垫/贴/腕垫 功放 显卡 游戏实物周边</td>
</tr>
<tr>
<td>1020</td>
<td>让你一步到位</td>
<td>笔记本玩游戏不方便</td>
<td>显示器 键盘 散热器/风扇 Hifi音箱 机箱 DIY兼容机 鼠标垫/贴/腕垫 功放 显卡 游戏实物周边</td>
</tr>
<tr>
<td>1021</td>
<td>真正的桌面PC</td>
<td>身临其境的游戏体验</td>
<td>显示器 键盘 散热器/风扇 Hifi音箱 机箱 DIY兼容机 鼠标垫/贴/腕垫 功放 显卡 游戏实物周边</td>
</tr>
<tr>
<td>146</td>
<td>让宝宝有趣味的玩耍</td>
<td>夏天也可以用的尿不湿</td>
<td>连身衣/爬服/哈衣 内衣套装 棉袄/棉服 睡袋/防踢被 纸尿裤/拉拉裤/纸尿片 学步鞋 三轮车 套装 餐椅 唐装</td>
</tr>
<tr>
<td>147</td>
<td>给宝宝更舒适的体验</td>
<td>婴儿的天然健康保护伞</td>
<td>连身衣/爬服/哈衣 内衣套装 棉袄/棉服 睡袋/防踢被 纸尿裤/拉拉裤/纸尿片 学步鞋 三轮车 套装 餐椅 唐装</td>
</tr>
<tr>
<td>148</td>
<td>宝宝冬季的小暖炉</td>
<td>为宝宝提供温暖的睡眠</td>
<td>连身衣/爬服/哈衣 内衣套装 棉袄/棉服 睡袋/防踢被 纸尿裤/拉拉裤/纸尿片 学步鞋 三轮车 套装 餐椅 唐装</td>
</tr>
</tbody>
</table>

**Baby’s small heater in winter**  **Provide warm sleep for your baby**
Cloud Theme Title Auto-Generation

Method2: generative model

Lace design is simple and elegant

Simple round neck long sleeve shirt
The lace design is simple and elegant,
Slim fit is more attractive,
...
<table>
<thead>
<tr>
<th>1</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>简约圆领长袖衬衫，蕾丝设计简约且优雅，修身的版型更为动人，3D印花更为休闲，很有文艺感。</td>
</tr>
<tr>
<td>3</td>
<td>眼下，外观小巧精美，看起来“萌萌哒”的多肉植物越来越受到养花爱好者喜爱，春天正是捣鼓这些可爱小植物的最佳时节。</td>
</tr>
<tr>
<td>4</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>5</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>6</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>7</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>8</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>9</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>10</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>11</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>12</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>13</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>14</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>15</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>16</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>17</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>18</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>19</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>20</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>21</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>22</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>23</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>24</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>25</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>26</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>27</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>28</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>29</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>30</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>31</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>32</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>33</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>34</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>35</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>36</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>37</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>38</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>39</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>40</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>41</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>42</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>43</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>44</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>45</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>46</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>47</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>48</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>49</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>50</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>51</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>52</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>53</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>54</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>55</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>56</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>57</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>58</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>59</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>60</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>61</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>62</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>63</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>64</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>65</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>66</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>67</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>68</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>69</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>70</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>71</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>72</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>73</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>74</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>75</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>76</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>77</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>78</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>79</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>80</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>81</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>82</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>83</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>84</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>85</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>86</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>87</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>88</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>89</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>90</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>91</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>92</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>93</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>94</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>95</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>96</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>97</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>98</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>99</td>
<td>蕾丝设计简约且优雅</td>
</tr>
<tr>
<td>100</td>
<td>蕾丝设计简约且优雅</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>生成title</th>
<th>生成title</th>
</tr>
</thead>
<tbody>
<tr>
<td>简约圆领长袖衬衫，蕾丝设计简约且优雅，修身的版型更为动人，3D印花更为休闲，很有文艺感。</td>
<td>多肉植物，给你最好的爱</td>
</tr>
<tr>
<td>眼下，外观小巧精美，看起来“萌萌哒”的多肉植物越来越受到养花爱好者喜爱，春天正是捣鼓这些可爱小植物的最佳时节。</td>
<td>蕾丝也要潮，穿出时尚新高度</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>简约白色，助你轻松成为街头焦点</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>高颜值水杯，喝水也要萌萌哒</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>穿上毛呢大衣，暖融融的新淑女</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>让你高大上的剃须体验</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>小清新连衣裙，怎么搭配都美</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>平价又好用，不输大牌的平价好物</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>时尚百搭的针织毛衣</td>
</tr>
<tr>
<td>蕾丝设计简约且优雅</td>
<td>实用家电，送爸妈更省心</td>
</tr>
</tbody>
</table>

炎炎夏日，小清新们福利又来了，这么美的你怎么少得了仙气爆棚的连衣裙？无论是文艺复古系长裙还是元气满满的短裙，搭配着简单的碎花或者结合民族风，都有着独特的情调，所以说小清新连衣裙是永不过时的美衣，选对连衣裙，怎么搭配都美。

奥黛丽赫本曾经说过：“不涂口红的女人没有未来”。由此可见，口红对于女人来说是至关重要的。而口红唇色也是特别重要的。由于唇色与年龄、肤色和肤色都有一定的关系。唇色偏红，偏暖，偏暗，偏冷，偏白，偏黑，偏粉，偏黄等，代表的气质特点也不同。
Cloud Theme Main Image Auto-Generation

1. Cloud theme cover image
2. Template Puzzle
3. TOP5/TOP7 Popularity selection from the cutout
4. Commodity Cutout
5. Detect and filter non-white background product images
6. Extract the TopN product map of each cloud-themed product
Step 1: Training data mining

Context-based query embedding & Query item bipartite Graph embedding
Cloud Theme Knowledge Card/Guide Generation

NLG model

PGN:

Soft template
Product embedding
Currently, items are retrieved through query cluster, keywords, question, and knowledge card;

Query cluster: Autumn suit, female student, fresh, college style...
Keywords: campus, in the Mood for Love, golden years
Question: Do you know the style of female students?
Knowledge card: We are at the border of immaturity and maturity. If we want to be a campus goddess, of course we have to choose what suits us. To become a dazzling scenery, we should live the most beautiful look. Come to the small society of university...
322工房纯羊绒打底衫 只需559元！

#省钱日#322工房6款纯羊绒打底衫，这个双12，只需要559你就可以把它带回家！

经典纯羊绒针织圆领V领打底衫，条纹短款轻薄针织羊绒衫女

原价799元 双12特价599元 满减到手价559元
Part 6: What is next?
Current Focus and Challenges

• Recommendation pipeline for Alibaba: billions of users and products with over trillions of edges and extremely rich attributes
  • Algorithm + System Co-optimization

• Graph embedding & inference for attributed heterogeneous network
  ✓ Heterogeneous nodes with attributes
  ✓ Different types of edges, easy to integrate knowledge (Knowledge Graph and Graph Embedding)
  ✓ For each specific task, best choice of edge type & node attributes

• Multi-modality & Scalable Bayesian Deep Learning

• Introducing high quality data labels
  ✓ High quality side information, such as high valuable users and merchandise
  ✓ Targeted positive and negative samples, node & edge importance weights

• Online Inference
  ✓ New data inflow, real-time graph update
  ✓ Can handle sequential information

• AliGraph Echo System On Cloud
AliGraph-Open Source

Welcome to Try out

SYSTEM+ALGORITHM

Code:
https://github.com/alibaba/graph-learn

Q/A :
graph-learn@list.alibaba-inc.com

Paper:
AliGraph: A Comprehensive Graph Neural Network Platform, VLDB 2019
Thanks

Xiaoyong Liu : xiaoyong.liu@alibaba-inc.com
Hongxia Yang : yang.yhx@alibaba-inc.com
Wei Lin : weilin.wl@alibaba-inc.com
Part Appendix : Cognitive
Merchandise Context

WeiTao

Product Title

Product details
A new style of colorful spring clothes. Korean style of temperament cashmere coat, medium and long models do not pick up the figure, 3D three-dimensional tailoring, button design unique, delicate patch, an indispensable one. With pointed square button shallow shoes, velvet The thick noodles are simply too much to put it down, and the colors are also very spring colors.
Short video front desk service

Guess you like the video recommendation

Taobao experience video content

Wow video
我比较喜欢穿背带裤，它卷到九分裤的位置

背带裤配短袖太幼稚，这样搭更性感
背带裤搭配短袖会显得非常孩子气，何不试试这样搭？
From the expert videos, the videos with higher products are selected as training data.

Generate a multi-product category data set with a single product image model.

Weak supervised object detection + Video description info.

Manual marking + few shot learning.
Video summary generation-1

- Video text description is the basic input information for video understanding

- Evaluate a large number of undesicriptive and unclear descriptions in the video
Video summary generation-2

- Video + seq2seq model
- Video content + review text content + product attributes, produce short video abstracts
  - Long sentences short
  - Long short sentences
  - Add Product Subject
- Manually filter the training data set from the videos with high views and high likes of users

Kid Toy Album

2019-03-22 颜色分类:豪华版 实测7岁以上 送送客

因为同学的推荐。又因为个人手机上玩过类似游
戏。果断入手。宝贝很欢喜。既能开发小脑，又可
以让其抛弃电子产品，提高专注力。杠杠滴！希望
用不了多久，宝贝破解技能超过老母。

很好玩！除了图纸造型外，孩子还搭了些其他造
型。店家发货速度快，质量跟专柜对比一样，应该
是正品！重点是真的很好玩！已经推荐给周围的朋友
了。

Develop the cerebellum to improve concentration toys

Variety of styles and counter quality
Video summary generation-3

- Youku: Film and television comprehensive video slices, automatically generate video tag.

- Training data: high-quality short video pool with complete titles
- Tags Auxiliary title generation
Personalized content themes

- Generate personalized content themes composed of products, graphics, and videos
- The video quality evaluation model automatically selects products from Youku short videos + hand-made short videos
- Gag recall, content-related recall
- Video sorting, video product mixing

人工智能生成性价比商品主题

需求场景
- 手淘消费内容不足，有近30%的搜索query无返回结果
- 消费内容人群特征弱，低端用户点击率低

01 商品集抽取
- 根据用户属性&query情况，个性化抽取商品，且可跨类目组合搭配商品集

方案优势
- 人工智能内容生成，可以针对长尾词定向生成内容，效度远高于人类生成。
- 推荐及智能写作可令内容更贴近人群

02 图集/视频生成
- 根据算法抽取的图集，筛选真实图片（如买家秀图）组成图集或集成视频，提升内容的置信度，提升消费者阅读兴趣