Tencent Ads: Interesting Problems and Unique Challenges

Haishan Liu
As the leading marketing platform in China, Tencent Ads has experienced a rapid growth over the recent years.

**2010-2018 Annual Revenue**

*Source: Financial report 2010-2018*

<table>
<thead>
<tr>
<th>Year</th>
<th>Tencent Corp</th>
<th>Tencent Ads</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>14</td>
<td>196</td>
</tr>
<tr>
<td>2011</td>
<td>20</td>
<td>285</td>
</tr>
<tr>
<td>2012</td>
<td>34</td>
<td>439</td>
</tr>
<tr>
<td>2013</td>
<td>50</td>
<td>604</td>
</tr>
<tr>
<td>2014</td>
<td>83</td>
<td>789</td>
</tr>
<tr>
<td>2015</td>
<td>175</td>
<td>1029</td>
</tr>
<tr>
<td>2016</td>
<td>270</td>
<td>1520</td>
</tr>
<tr>
<td>2017</td>
<td>404</td>
<td>2378</td>
</tr>
<tr>
<td>2018</td>
<td>581</td>
<td>3127</td>
</tr>
</tbody>
</table>

8.24B USD
Tencent Marketing Solution Core Capability – Panorama Connection

Top Internet Traffic Sites in China
Tencent Marketing Solution Core Capability – Digital Intelligence

Covers 1B+ Chinese internet users

source: Penguin 2018 Data
Tencent 2019年Q1 Financial Report
Tencent Marketing Solution Core Capability – Diverse & Engaging Formats

- Xiaomi x Moment
- Lancôme x QQ
- Cadillac x News Feed
- Nike x QQ Music
- McDonald x WeChat Game

- Celebrities in Moment
- AI Cosmetic Try-up
- Commercial Content
- Runner’s Radio
- Gamelet
Tencent Marketing Solution Core Capability – Diverse & Engaging Formats

Enhanced Branding via Social Endorsement

WeChat Moment @Friend

QQ Red Packet
Tencent Ad Ecosystem Helps Businesses Monetize Site Traffic

**优量汇**
Serving ads on Tencent Audience Network with the help of Tencent Ad Platform technologies

<table>
<thead>
<tr>
<th>No. of Quality APPs</th>
<th>Hero APP %</th>
</tr>
</thead>
<tbody>
<tr>
<td>100k+</td>
<td>75%</td>
</tr>
<tr>
<td>Mobile MAU</td>
<td>100M+</td>
</tr>
<tr>
<td>500M</td>
<td></td>
</tr>
</tbody>
</table>

**微信广告**
Certain official account and applet owners can monetize their content by utilizing the built-in ad slots

- 3B+  
  First 3Q of 2018 Shared Revenue

- 600+  
  No. of site owners with 100k revenue/month

- 40%  
  No. of individual site owners

- 40%  
  Revenue for individual site owners

**Monetization for WeChat Official Accounts**
- Article Bottom
- Article Middle
- Reciprocal
- Cash-back products

**Monetization for WeChat Applets**
- Banner
- Incentive video
- Splash screen
- Reciprocal
Tencent Ad: Technology for Good

Charity Advertising Competition

2017-2018 Charity Advertising Facts:

- **160M+** In funding
- **1700+** Participating teams
- **62** Award winning pieces
- **1,000M+** People reached for engagement

---

**《一个人的球队》**
直击器官捐献话题
小屏感动大屏
由社交平台发酵的社会话题

**《灯山行动》**
用微信小游戏体验上学路
化氪金为公益

**《不朽的丰碑》**
以本地推广广告唤起共鸣
致敬抗战老兵

**《一起捐脸，勇敢代言》**
联动AI，天天P图
让每个人为唇腭裂发声代言

**《一个人的球队》**
直击器官捐献话题
小屏感动大屏
由社交平台发酵的社会话题
Machine Learning Applications

Smart Targeting

Smart Creative

Smart Bidding
### Example Analysis of a Free-Text Chinese Query

**Question:** 喜剧脱口秀节目《吐槽大会》为什么能火？

<table>
<thead>
<tr>
<th>Analysis Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segmentation:</strong></td>
<td>喜剧/脱口秀/节目/《/吐槽/大会/》/为什么/能/火/?/</td>
</tr>
<tr>
<td><strong>PoS tagging:</strong></td>
<td>喜剧/n 脱口秀/n 节目/n 《/w 吐槽/v 大会/n 》/w 为什么/r 能/v 火/n ?/w</td>
</tr>
<tr>
<td><strong>NER:</strong></td>
<td>喜剧/n 脱口秀/n 节目/n 《/w 吐槽大会/NS 》/w 为什么/r 能/v 火/n ?/w</td>
</tr>
<tr>
<td><strong>BoW:</strong></td>
<td>喜剧:0.45, 脱口秀:0.52, 节目:0.36, 吐槽:0.39, 大会:0.38, 火:0.31</td>
</tr>
<tr>
<td><strong>Keywords extraction:</strong></td>
<td>吐槽大会:0.92, 喜剧:0.78, 脱口秀:0.88</td>
</tr>
<tr>
<td><strong>Keywords expansion:</strong></td>
<td>情景喜剧:0.84, 喜剧电影:0.80, 今晚80后:0.84, 相声:0.74</td>
</tr>
<tr>
<td><strong>LDA:</strong></td>
<td>8867:0.15, 明星(0.09) 中国(0.033) 韩国(0.03) 歌手(0.03) 香港(0.02) 演员(0.02) 8191:0.11, 赵本山(0.05) 小品(0.03) 春晚(0.02) 小沈阳(0.01) 喜剧(0.01) 台词(0.01)</td>
</tr>
<tr>
<td><strong>Embedding:</strong></td>
<td>1:0.32, 2:0.56, 3:0.01, 4:0.66, 5:-0.12, 6:0.89, 7:-0.45, 8:0.23, 9:0.14, 10:0.54</td>
</tr>
<tr>
<td><strong>Classification:</strong></td>
<td>06:娱乐休闲—综艺—喜剧:0.95, 05:娱乐休闲—综艺—脱口秀:0.93</td>
</tr>
</tbody>
</table>
NLP at Tencent Ads – an Overview

Service
- TextMiner RPC Service (C++/Python/Java/Scala)
- TextMiner Lib (C++)

Module
- Text Classification
  - BoW
  - Keyword extraction
  - Keyword expansion
  - LDA
- Chinese Segmentation
- PoS Tagging
- NER
- Embedding

Tool
- TensorFlow
- PyTorch
- FastText
- FAISS
- MaxEnt
- Distributed LDA: Peacock/LightLDA
- Distributed Embedding: AnyEmbedding
- DNN + NLP: DeepText
- Crawler (TEG)
- Data Kitchen (MIG)
- XGBoost/LightGBM
- Labeling: LabelMe
- Distributed LDA: Peacock/LightLDA
- Distributed Embedding: AnyEmbedding

Data
- Multi-context behavioral data
- Dictionary
- Industry Knowledge Repo
- Word Bank
AnyEmbedding – A Unified Framework for NLP

- Generate embeddings for any entities of interest
- A unified framework to tackle various NLP problems
AnyEmbedding – A Distributed Embedding System

- Goal: Design arbitrary models and express arbitrary entities via “connections”
  - word2vec, sent2vec, doc2vec, FastText, Hierarchical Document Vector model
  - TransE, TransR
  - node2vec
- Flexible mode of connections
  - 1-to-1: special case of 1-to-many
  - 1-to-many: SkipGram
  - Many-to-1: CBOW, Concat
  - Many-to-many: extended many-to-1
- Supports both single-node and distributed computation
  - Scales up to 100k words & 1b users
  - Async distribution
- Optimization method
  - supports Negative Sampling (NCE loss)
SkipGram Negative Sampling (SGNS)

**Objective Function**

\[
L = \sum_{i,j \in B_h} \sum_{k \neq j, k \leq b_{i,j}} \left[ \log \left( \frac{\sigma(u(w_{i,j})v^T(w_{i,k}))}{\sigma(u(w_{i,j})v^T(w_{i,j}))} \right) + \sum_{\hat{w} \in N_{i,j,k}} \log(1 - \frac{\sigma(u(w_{i,j})v^T(\hat{w}))}{\sigma(u(w_{i,j})v^T(w_{i,j}))}) \right]
\]

**Gradient**

\[
\nabla L | u(w_{i,o,j}) = \sum_{\{i,j\} \in B_h} \sum_{k \neq j, k \leq b_{i,j}} \left[ \frac{(1 - \sigma(u(w_{i,j})v^T(w_{i,k})))v(w_{i,k})}{\text{dot product}} - \sum_{\hat{w} \in N_{i,j,k}} \sigma(u(w_{i,j})v^T(\hat{w}))v(\hat{w}) \right] \]

- \(V\): Vocab Dict
- \(B_h\): Set of vocab in a batch
- \(\sigma\): sigmoid function
- \(w_{i,j}\): words in \(B_h\)
- \(b_{i,j}\): sliding window size
- \(w_{i,k}\): words in the context of \(w_{i,j}\)
- \(\hat{w}\): negative sampling for \(w_{i,j}\)
- \(N_{i,j,k}\): set of all negative samples
- \(u\): input embedding
- \(v\): output embedding

Copyright © 1998 - 2018 Tencent Inc. All Rights Reserved
Distributed Computation

- Parameter Server Group
  - Each PS stores embedding shards
  - Receives minibatches from Worker Group, calculates the dot product with the context, and sends back to Worker Group
  - Receives the result of linear combination from Worker Group and update parameters

- Worker Group
  - Loads minibatches, and sends to PS
  - Receives computed dot products from the PS Group, computes the linear combination and sends to PS

- Network transmission
  - Minibatch data, integer vectors
  - Dot product, Linear combination, real-valued scalars

Ordentlich, E.; Yang, L.; Feng, A.; Cnudde, P.; Grbovic, M.; Djuric, N.; Radosavljevic, V.; and Owens, G. Network-efficient distributed word2vec training system for large vocabularies. CIKM 2016.
## DeepText: An Open Platform for Deep NLP

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Model</th>
<th>Common methods</th>
<th>Example tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text Classification</strong></td>
<td>Classify input texts to predefined categories</td>
<td>s → c</td>
<td>SVM, MaxEnt, TextCNN, TextRNN</td>
<td>Text classification, sentiment analysis, intent understanding</td>
</tr>
<tr>
<td><strong>Structural prediction</strong></td>
<td>Derive structured information from free text</td>
<td>s → [s]</td>
<td>HMM, MEMM, CRF, Structure Perceptron, LSTM-CRF, CNN-LSTM-CRF, Transition-based</td>
<td>Chinese segmentation, PoS tagging, NER, syntax analysis</td>
</tr>
<tr>
<td><strong>Text Matching</strong></td>
<td>Match texts based on various criteria</td>
<td>s, t → R⁺</td>
<td>RankNet, GBRank, LambdaMART, DSSM, CDSSM, MatchPyramid</td>
<td>Information retrieval, Search engine, QA bot</td>
</tr>
<tr>
<td><strong>Text Transformation</strong></td>
<td>Convert a input sequence to an output sequence</td>
<td>s → t</td>
<td>SMT, NMT (Seq2seq+Attention)</td>
<td>Machine translation, QA bot, text generation</td>
</tr>
</tbody>
</table>

**Goal:** To develop an open NLP experiment platform to tackle the problem of text classification, tagging, and matching, and support a wide range of models and optimization methods proposed in the field for fast proof of concept and experimentation.
DeepText-Classification: a Wide Range of Supported Methods

**Classifier Layer**
- Binary Class
- Multi Class
- Multi Label
- Hierarchical Class
  - Softmax Cross Entropy
  - Sigmod Cross Entropy
  - NCE Loss

**Model Layer**
- MLP
- FastText
- TextCNN
- LSTM/GRU
- Bi-LSTM/GRU
- AttentiveConvNet
- ... (Other methods)
  - Attention
  - Shortcut/Highway
  - Dropout
  - Normalization
  - ...

**Embedding Layer**
- Pretrained embedding
- Region embedding
- embedding_lookup
- embedding_lookup_sparse
- Custom Feature

**Input Layer**
- Encoding (SUM/AVG/CNN/LSTM)
- Token
- Token-Ngram
- Char
- Char-Ngram
- Custom Feature
  - Pic(encoding)
  - LDA Topic
  - Keywords
  - ...

*Copyright © 1998 - 2018 Tencent Inc. All Rights Reserved*
DeepText-Matching: a Variety of Flexible Matching Models

Matching Layer
(representation-based /interaction-based)

Representation layer
(BoW/CNN/RNN)

Input layer
(word2vec/GloVe/Sent2vec)

k: keyword
d: doc

Matching score

Representation-based Match (cos)
Representation-based Match (MLP)

Shared lookup table

Unified framework for deep semantic matching: MatchNet

Interaction-based Match

Conv
Pooling

word interaction tensor
n-gram interaction tensor
User Profile Understanding & Accurate Audience Insights

Demographics
- inherit
- location
- education
- marriage
- wealth
- employment

Device
- carrier
- network
- brand
- type
- OS
- price

Interest
- purchasing intent
- hobby

Behavior
- travel
- App
- eCommerce
- O2O
- Ad engagement
- purchase

Vertical
- auto
- education
- eCommerce
- game
- travel
- 3C
- finance
- FMCG

Custom
- seed
- expansion
- 1-party labels

UID
Traditional Audience Definition - Manual Targeting

Advertisers manually specify the targeting criteria. Boolean match during ads retrieval.

Disadvantages:
- Domain knowledge
- Information loss
- Low granularity
- Manual optimization

Large number of targetable attributes

Gender, age, geo, education.
Marriage, employment, wealth, culture.
Context, device, weather.

Interests, keywords.
Followers, APP installs, payments, engagements.
APP, eCommerce, Search, and Social behaviors.

Auto Targeting: Automatic Audience Identification by AI

Multiple ways to define the seed audience

Targeting by machine learning models

Matching based on relevance

A change of mindset
• Targeting by user-ad relevance

- Existing leads from 1st-party data
- Previously converted users
- User/Ad features

- Matching based on relevance

- Targeting by user-ad relevance

- Auto Targeting: Automatic Audience Identification by AI
Auto-Targeting: Representation Learning & ANN

**Representation Learning**

Learns embeddings for users and ads; can be used during ad retrieval

- User embedding
- Ad embedding

**Approximate Nearest Neighbor Search**

Efficient ad retrieval based on user/ad embeddings under scalability constraints

Lookalike Expansion for Targeting: Identify Potential New Customers

**Custom audience based on known converters**
- 1st-party data
- Tencent data

**User Profiles**
- Mined user features

**Lookalike**
- Understand the commonality of the seed audience
- Find lookalike audience

**Peacock**
- Large scale distributed LDA

**Logistic Regression**
- One model per seed audience, expansion done offline

**DNN**
- Deep embedding for each seed audience, expansion done online

2013
2014
2015
AI in Tencent Social Ads

Smart Targeting  Smart Creative  Smart Bidding
Manual Creative Review: Laborious, Inefficient, and Error-Prone

The manual review procedure consists of **Hundreds** of individual rules

1. **Text** - suggestive, vulgar, over the top, low effort
2. **Image** - blurry, deformed, IP violation, % of text
3. **Landing page** - relevance to the creative content
4. **Official Account** - title, description
5. **Published content** - history, compliance

More than 4x increase of the review request in 1 year

1. **Hard to scale with human labor**
   - Extremely high volume, Repeated ad submission
2. **Steep learning curve, prone to inconsistency**
   - Complicated set of rules, high training cost, implicit knowledge that is hard to be passed down
3. **Some cases are hard for human**
   - Celebrity/IP/Plagiarism detection
Low-Quality Content Detection based on Texts or Images Individually

1. Text-based detection

“是男人就来野战一场，一小时的持久战斗！”
Be a man and do a field battle, an hour-long game of stamina

Review result: Reject
Reason: sexual innuendo

2. Image-based detection

Goal is not to reject... but to help advertisers improve

Review result: Reject
Reason: explicit/suggestive content
Low-Quality Content Detection based on Joint Image/Text model

3. Joint image/text models

"Finally you are here. Let me give you some warmth."

10+ similar models
Expert System for Rejection

1. Expert system to reason about the cause of rejection
   - LCS = 0.85
   - Review result: Rejection
   - Cause of rejection: suggestive content in context
   - LCS <= 0.7

2. Auto approval
   - Precision: 94.0% 93.4% 100.0%
   - Recall: 60.0%
   - Manual: 100.0%
   - CNN+Wide: 60.0%
   - Cross feature learning
   - Auto approval decision
   - Text Embedding
   - convolution pooling
   - Serving history sparse: ID+industry
   - Rule engine Template engine
   - Rule DB Cause DB
   - Rule Configuration System

If no applicable rule
VideoIn Ads: Smart Product Placement

Ad Slot Detection
Find suitable frames for placement

Ad Creative Rendering
Render the asset with the ambient brightness, saturation, etc.

Object & Asset Orientation
Adjust the perspective of 3D models and assets

Video Tracking
Position the rendered creative in the frames
Exclude unsuitable sections by detecting camera cuts from a grouping of key frames
Detection of the first frame of a potential ad slot

Non-central object masking algorithm based on object segmentation with mask R-CNN

Potential Ad Area Location
Detection of the first frame of a potential ad slot

Color clustering and bounding box search to detect empty table surface

Potential Ad Area Location
Picture frame detection and tracking

In-house labeling of 17 classes of video contents

Object recognition and segmentation
Based on DaSiamRPN
Typical Ad Format

- User supplied pictures
- User supplied pictures & platform supplied 3D assets
- User supplied 360-degree pictures
Detecting the bottom edge of the table corner

Matching the bottom edge of the object
In: properly oriented stock picture

Out: adjusted picture with background harmonization
1. Original
2. Brightness
3. Saturation
4. Blurriness
5. Shadow
Pre-adjustment

Post-adjustment
青檀姐
Static Creative Is Hard To Optimize For Advertisers with A Large Inventory

Static Creative: 

- Hit-or-miss
- Coarse-grained targeting
- Static content with no personalization
DPA: Automatic creation, retrieval, and rendering of personalized Ad Creatives

- **Advantages**
  - **Automatic creative generation**
    - No need for trial-and-error optimization from advertisers
  - **Fine-grained targeting**
    - 100m targetable products
    - More detailed and accurate than ad targeting
  - **Truly personalized**
    - Different experience even for the same ad

**Process:**
- **Creation**
- **Serving**
- **Impression**

**Flow:**
- **Advertiser**
- **User**

**Components:**
- **Product DB**
- **Product targeting**
- **Online creative generation**
- **Automatic creative generation**
Anatomy of a Creative

- **Efficiency**: Automatic creative generation based on product + template anatomy

- **Product**: Programmatic interface to sync from advertisers’ inventory

- **Template**: Coverage over more than 3000 ad slots, automatically generated by AI

- **Performance**: optimize \( \argmax \{ p_{CTR}(\text{Product}, \text{Template} | \text{User}) \} \)
**Layout**: Defines the type, position, and size of template elements

**Element**: Individual graphic components, e.g., background, decoration, textual decorations

10k+ Layers x 700k+ Elms = ?
Template Generation: Embedding for Element Combination

- Treat element selection as a combinatorial optimization
- Map elements through a Deep Attention Network

Result: Generated 5.5M+ templates. Manual evaluation from sampling shows 73% achieve designer grade
Product Retrieval based on Learned User and Product Representations

Training

Click/browse data

Product embeddings

Hidden layer

Image embedding
Category embedding
Name embedding

Product image
Product category
Product name

CNN
Text-CNN

User

Browsed product embedding

Geo embedding
App embedding
Gender embedding
Keyword embedding

User embedding

History browsing data

ANN

Product index

Prediction

Product embedding 1
Product embedding 2
Product embedding 3
Demo of Generated Templates
Machine Learning Applications

Smart Targeting  Smart Creative  Smart Bidding
Optimized Cost Per Acquisition (oCPA) Bidding

- Accurate prediction of pCVR is the key to acquiring the most conversions within cost and budget.

- **Engagement Funnel**
  - Impression
  - Click
  - Download
  - Activation
  - pay

Conversions (Acquisitions)

- oCPA lets the advertisers specify a target conversion type to optimize.
- Advertisers specify the target cost per conversion.
- The system adjusts the bid on behalf of advertisers based on the predicted conversion rate (pCVR).
Predicting CVR Accurately is the Key to oCPA Bidding

- Tracking log
- User profile
- Landing page
- Merchandise data
- Conversion data

Large-scale parallel training system

- SmartBid
- pCVR
- pCTR

Impression → Conversion

Continuous feedback loop to improve pCVR

Predicting CVR Accurately is the Key to oCPA Bidding
Large-Scale Parallel Training Framework

P2P topology
Parameter servers and workers are collocated
Balance different load profiles of the two roles
**Scalability Test: Processing Power**

**Model configuration**
- Feature dimension: 5
- Input: 890
- 2 hidden layers: 135, 67 nodes
- Output: 1 node

**Observation**
- Linear increase in processing power with respect to number of machines
- Successfully trained on 100B features
Scalability Test: Speedup

**Model structure:**
- Feature vector dim: 11
- Input: 110
- 5 hidden layers: 400, 400, 300, 200, 100
- Output: 1

**Data:**
- Training example: 6B
- Feature dimension: 4.8B

**Observation:** Linear acceleration wrt the number of machines

**Speedup vs number of nodes**

- Speedup vs Number of machines
- Training example: 6B
- Feature dimension: 4.8B

- Model structure:
  - Feature vector dim: 11
  - Input: 110
  - 5 hidden layers: 400, 400, 300, 200, 100
  - Output: 1
Deep conversion objectives have a very long delay

CDF of an example conversion objective

Delay model

Conversion model

Joint training of conversion and delay model

Click download link
Download in app store
Open app (activation)
Register in app
Advertiser report
激活数据按点击归因
微信广告平台提取转化数据
Thank you!