

# 基于注意力卷积的神经机器翻译

汪 琪 段湘煜

(苏州大学计算机科学与技术学院 江苏 苏州 215006)

**摘 要** 现有神经机器翻译模型普遍采用的注意力机制是基于单词级别的,文中通过在注意力机制上执行多层卷积,从而将注意力机制从基于单词的级别提高到基于短语的级别。经过卷积操作后的注意力信息将愈加明显地体现出短语结构性,并被用于生成新的上下文向量,从而将新生成的上下文向量融入到神经机器翻译框架中。在大规模的中-英测试数据集上的实验结果表明,基于注意力卷积的神经机器翻译模型能够很好地捕获语句中的短语结构信息,增强翻译词前后的上下文依赖关系,优化上下文向量,提高机器翻译的性能。

**关键词** 神经机器翻译,多层卷积网络结构,注意力机制,短语级别

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.11.035

## Neural Machine Translation Based on Attention Convolution

WANG Qi DUAN Xiang-yu

(School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu 215006, China)

**Abstract** The attention mechanism commonly used by the existing neural machine translation is based on the word level. By creating multi-layer convolutional structure on the basis of attention mechanism, this paper improved attention mechanism from word-based level to phrase-based level. After convolutional operation, the attention information can reflect phrase structure more clearly and generate new context vectors. Then, the new context vectors are used to integrate into the neural machine translation framework. Experimental results on large-scale Chinese-to-English tasks show that neural machine translation based on attention convolution can effectively capture the phrasal information in statements, enhance the context dependencies of translated words, optimize the context vectors and improve the translation quality.

**Keywords** Neural machine translation, Multi-layer convolutional structure, Attention mechanism, Phrase-based level

## 1 引言

机器翻译(Machine Translation, MT)是指借由计算机将一种人类语言自动翻译为另一种语言的自动翻译过程<sup>[1-3]</sup>。传统的机器翻译方法有基于统计的机器翻译<sup>[4]</sup>和基于实例的机器翻译。近年来,随着神经网络的发展,利用神经网络将源语言映射到目标语言的翻译模型,即神经机器翻译模型(Neural Machine Translation, NMT)<sup>[5-8]</sup>,显著提升了机器翻译的质量,超越了传统机器翻译方法的性能,成为目前工业界和学术界的主流方法。

NMT 系统使用了编码-解码(Encoder-Decoder)框架,将源语言通过编码器编码成隐藏向量信息,再通过解码器将该隐藏向量信息解码成目标语言,编码器和解码器之间通过注意力机制(Attention Mechanism)连接,在翻译目标单词时检测其与源端语句相关的部分<sup>[5]</sup>。

现有 NMT 系统的注意力机制将注意力放在源语句的各个单词之上,而本文提出的新的注意力机制将注意力放在源

语句的短语之上。由于卷积网络可以学习局部信息<sup>[9-10]</sup>,所提方法通过在注意力的基础上使用多层卷积网络,学习获得局部注意力,即短语级别的注意力。该方法有利于捕捉翻译过程中的短语级别的上下文依赖关系。实验结果表明,相比于基准 NMT 系统,本文所提方法能有效提高机器翻译的质量。

## 2 相关工作

近年来,许多研究学者对 NMT 框架中的注意力机制和神经网络模型展开了多种研究和分析,推动了机器翻译的快速发展。

2015 年, Luong 等<sup>[6]</sup>在 Bahdanau 等<sup>[5]</sup>研究的基础上对注意力机制的计算问题提出了各种优化方案,并且提出了一种局部的注意力机制(Local Attention),使得解码端生成目标单词时着重考虑源端某一部分的隐藏层信息; Cheng 等<sup>[11]</sup>提出使用联合训练的双向注意力模型,以获得更好的注意力信息; 2016 年, Tu 等<sup>[12]</sup>在注意力机制的基础上提出了 Coverage

投稿日期:2018-04-18 返修日期:2018-07-12 本文受国家自然科学基金(61673289),国家重点研发计划“政府间国际科技创新合作”重点专项(2016YFE0132100)资助。

汪 琪(1994-),女,硕士生,CCF 会员,主要研究方向为自然语言处理、机器翻译, E-mail: littlewqq@gmail.com; 段湘煜(1976-),男,副教授,主要研究方向为自然语言处理、机器翻译, E-mail: xiangyuduan@suda.edu.cn(通信作者)。

方法来解决机器翻译中“过翻译”和“欠翻译”的问题;2017年,Gehring 等<sup>[9]</sup>提出了一种完全基于卷积神经网络的模型;同年,Vaswani 等<sup>[13]</sup>提出了一种完全使用注意力机制来实现翻译的模型。

区别于以上提及的在注意力机制上进行的各种优化和改进工作<sup>[6,11-13]</sup>,本文方法在现有的注意力机制的基础上添加卷积网络结构,利用卷积网络可以学习语句局部信息的特点,在经过多层卷积之后可以学习得到更大范围的局部信息,从而获得高层次的注意力信息,以优化上下文向量。

Gehring 等<sup>[9]</sup>提出的模型使用卷积神经网络(Convolutional Neural Network,CNN)替代编码端和解码端的循环神经网络(RNN)。与 RNN 的链式结构相比,CNN 采用的是层级结构,CNN 框架通过对源端和目标端的语句序列进行多层卷积运算,加快了运算速度,同时也对语句结构中的复杂关系信息进行了更好的处理和提取。但是这种方法可能会导致源端和目标端短语对齐不一致的情况,因为编码端和解码端的两个 CNN 结构是互相独立的,当源端切分成  $N$  个短语时,目标端可能被切分成  $M(N \neq M)$  个短语,出现对齐效果差的情况。

针对这种缺陷,本文提出基于注意力卷积的神经机器翻译,在注意力机制的基础上结合卷积网络结构,利用多层的卷积网络模型获得短语级别的注意力信息;同时,本文在编码端和解码端使用链式的 RNN 结构,避免短语切分不一致的情况。

### 3 基准 NMT 系统

神经机器翻译使用端到端的翻译模型<sup>[5,7]</sup>,它在给定源端语句  $X = x_1, x_2, \dots, x_{|X|}$  的条件下直接对目标端语句  $Y = y_1, y_2, \dots, y_{|Y|}$  的条件概率  $P(Y|X)$  进行建模。编码端和解码端均使用 RNN 对语句进行编码建模,在实际应用中,编码器将源端语句编码为源端隐藏层向量  $H = h_1, h_2, \dots, h_{|X|}$  时通常使用门循环单元(Gated Recurrent Unit,GRU)<sup>[14]</sup>或者长短期记忆网络(Long Short-Term Memory,LSTM)<sup>[15]</sup>。

$$h_i = f(x_i, h_{i-1}) \quad (1)$$

其中,  $x_i$  是  $i$  时刻的源端单词,  $h_{i-1}$  是  $i-1$  时刻的源端隐藏层向量,  $f$  是 GRU 或 LSTM 方法。解码器则根据条件概率  $P(y_i)$  逐字地预测出目标语句  $Y$ 。在生成  $Y$  时,解码器利用目标端的隐藏层向量信息、目标端前一时刻的单词信息以及通过注意力机制得到的上下文向量计算目标端单词的条件概率:

$$p(y_j | y_{<j}, X) = g(s_j, y_{j-1}, c_j) \quad (2)$$

其中,  $g$  是一个非线性函数,  $y_j$  是目标端  $j$  时刻的单词,  $s_j$  是  $j$  时刻的隐藏层向量,  $c_j$  是  $j$  时刻的上下文向量,计算公式如下:

$$s_j = f(s_{j-1}, y_{j-1}, c_j) \quad (3)$$

$$c_j = \sum_{i=1}^{|X|} \alpha_{ji} h_i \quad (4)$$

$$\alpha_{ji} = \frac{\exp(e_{ji})}{\sum_{k=1}^{|X|} \exp(e_{jk})} \quad (5)$$

$$e_{ji} = \vartheta_a^T \tanh(W_a s_{j-1} + U_a h_i) \quad (6)$$

其中,  $\vartheta_a^T, W_a, U_a$  是参数矩阵,  $\alpha_{ji}$  是权重信息。通过这种注意

力机制可以在生成目标端单词时构建当前词对源端所有词的注意力分布。

## 4 基于注意力卷积的神经机器翻译

在现有的神经网络机器翻译研究的基础上,本文提出使用基于注意力卷积的神经机器翻译。通过建立多层卷积网络结构,丰富源端和目标端语句中相对应短语的注意力信息,抓取更丰富的结构信息为机器翻译提供更好的帮助。

所提方法的基本框架如图 1 所示。

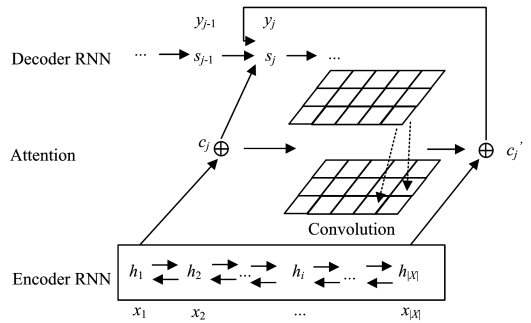


图 1 基于注意力卷积的神经机器翻译的基本框架

Fig. 1 Basic framework of attention convolution based on NMT

### 4.1 卷积结构

在图像处理中,卷积操作是指使用一个卷积核对图像中的每个像素点进行一系列的操作<sup>[10]</sup>。本文对注意力信息建立多层卷积网络结构,丰富语句中前后单词的依赖信息,再通过门控机制对信息流进行控制和更新。

在神经机器翻译中,注意力信息包含求解每一时刻目标端单词所对应的源端语句单词的概率分布。假设源端(src)语句为  $\{x_1, x_2, x_3, x_4, x_5, x_6\}$ , 目标端(trg)语句为  $\{y_1, y_2, y_3, y_4, y_5\}$ , 如图 2 所示,灰度图颜色越深的位置注意力权重越大。

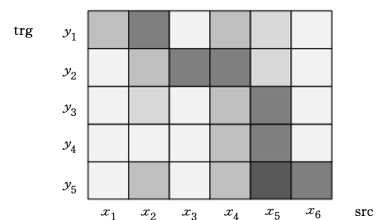


图 2 注意力信息的灰度示意图

Fig. 2 Grayscale diagram of attention information

因此,为了提高机器翻译的效果,本文提出基于注意力卷积的方法来增强机器翻译过程中注意力信息前后的短语关联性,通过层级(例如 6 层、10 层等)的卷积网络结构,使得高层的卷积网络包含短语块的信息数据,以增强语句上下文之间的依赖关系。

假设卷积核大小为  $2 * 1$ , 源端语句为  $\{x_1, x_2, x_3, x_4, x_5, x_6\}$ , 对目标端当前时刻的注意力信息建立多层卷积结构,如图 3 所示。可以看出,初始状态下,  $\{x_1\}$  或  $\{x_4, x_5\}$  为当前时刻目标端单词对应的词组的概率较大,随着卷积层数的增加,高层的信息分布包含了低层的信息,逐渐形成了一种层级的注意力信息分布。

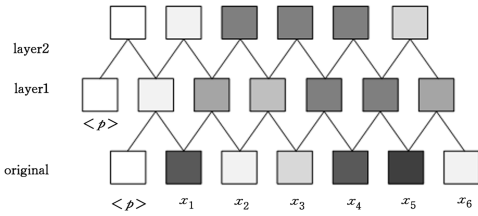


图3 对当前时刻注意力建立的多层卷积结构

Fig. 3 Multi-layer convolutional structure based on attention at present

#### 4.2 门控机制

为了使网络具有控制信息流传递的能力,在RNN中通常会引入LSTM或者GRU单元<sup>[7,13-16]</sup>,因此,本文对多层卷积网络的输出信息增加了一个门控机制,用以控制输出信息的传递。

新增的门控机制如图4所示。将经过多层卷积结构网络的注意力信息投入到门控机制中,再结合sigmoid和softmax函数,对其进行限制和优化。

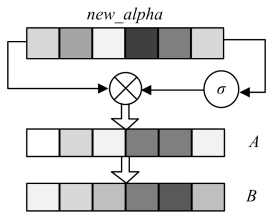


图4 新增的门控机制

Fig. 4 Additional gating mechanism

新增的门控机制函数如下:

$$A = \text{new\_alpha} \otimes \sigma(\text{new\_alpha}) \quad (7)$$

$$B = \text{softmax}(A) \quad (8)$$

其中,  $\text{new\_alpha}$  是新的注意力信息,即通过多层卷积网络结构得到的卷积函数激活值。sigmoid函数将  $\text{new\_alpha}$  映射到(0, 1)的区间,再利用点积对两者相乘,得到门控机制过滤后的分布值  $A$ ,最后再将其输入到softmax函数中,得到最终的注意力信息  $B$ 。

#### 4.3 融合的注意力机制

得到新的上下文信息后,我们利用如下公式将其融入到原有框架中:

$$c_j' = \sum_{i=1}^{|X|} \alpha_{ji}' h_i \quad (9)$$

$$p(y_j | y_{<j, X}) = g(s_j, y_{j-1}, c_j, c_j') \quad (10)$$

其中,  $\alpha_{ji}'$  是新的注意力分布信息,  $h_i$  是源端的隐藏层向量,  $c_j'$  是新得到的上下文向量。将原有的上下文向量  $c_j$  与  $c_j'$  直接相连接,通过上下文向量的融合,使得模型在解码时能融入更丰富的上下文信息,进而产生较好的译文。

### 5 基于注意力卷积的机器翻译训练与实验结果分析

#### 5.1 实验设置

在对注意力层建立层级卷积结构的过程中,为了研究不同卷积形状和网络深度环境下得到的训练模型的性能差异,

本文在使用同一语料集的情况下,设置不同的多层卷积网络结构进行对比实验。

本文实验采用的是DyNet2.0系统<sup>[17]</sup>和相应的Lamtram代码<sup>[18]</sup>。使用的训练集是125万句的LDC(Linguistic Data Consortium)中-英平行双语句对,测试集和开发集均采用美国国家标准与技术研究院2002年的数据,其中NIST02, NIST03, NIST04, NIST05和NIST08作为测试集, NIST06作为开发集。实验中,设置LDC语料中的前3万个高频词,其余用<unk>代替,设置学习率为0.001, Dropout为0.4;解码采用集束搜索(Beam Search)设置为5的情况,其余采用默认设置。首先使用基准的Lamtram系统实现神经机器翻译,得到基准模型;然后将基准模型读入提出的框架中,固定基准模型中的参数,增加多层卷积网络的相关参数,得到基于注意力卷积的翻译模型。

为了评价模型的翻译效果,本文使用不区分大小的BLEU-4作为系统翻译质量的评测指标,脚本为multi-bleu.perl<sup>1)</sup>。BLEU<sup>[19]</sup>是一种国际通用的机器翻译评测脚本,利用语句中的  $n$  元词组的精确度和句长的惩罚因子将机器翻译产生的译文和专职的人工翻译译文进行比较,计算出BLEU值。BLEU值的范围是0~100,数值越高表明机器翻译的效果越好。

在建立多层卷积网络结构的过程中,我们主要考虑卷积核大小和网络深度的设置。考虑到二元词组和三元词组的常用性,实验分别采用大小为  $2 \times 1$  和  $3 \times 1$  的卷积核,测试网络深度为3, 6, 10, 12层。

#### 5.2 实验结果及分析

##### 5.2.1 实验结果

本节主要进行两组对比实验,每组再分别进行多个对比实验,用以验证对于不同的参数配置,所提模型能否提升译文质量。实验结果如表1、表2所列。

表1 卷积核大小为  $2 \times 1$  时在LDC数据集上的BLEU值Table 1 BLEU values when kernel size is  $2 \times 1$  on LDC data set

实验系统	NIST02	NIST03	NIST04	NIST05	NIST08	AVG.	
基准系统	37.85	34.81	38.64	35.17	26.41	34.576	-
+3层网络结构	39.02	35.68	39.58	36.02	27.14	35.488	+0.912
+6层网络结构	38.34	36.29	39.31	36.56	27.17	35.534	+0.958
+10层网络结构	38.75	35.87	39.62	36.33	26.99	35.512	+0.936
+12层网络结构	38.48	35.91	39.18	36.05	26.79	35.282	+0.706

表2 卷积核大小为  $3 \times 1$  时在LDC数据集上的BLEU值Table 2 BLEU values when kernel size is  $3 \times 1$  on LDC data set

实验系统	NIST02	NIST03	NIST04	NIST05	NIST08	AVG.	
基准系统	37.85	34.81	38.64	35.17	26.41	34.576	-
+6层网络结构	38.50	35.77	39.62	36.38	27.51	35.556	+0.98
+10层网络结构	38.76	35.81	39.31	36.30	26.90	35.416	+0.84
+12层网络结构	38.68	36.12	39.44	36.08	26.72	35.408	+0.832

<sup>1)</sup> <https://github.com/moses-smt/mosesdecoder/blob/master/scripts/generic/multi-bleu.perl>

表 1 中,设置卷积核大小为  $2 \times 1$ ,多层卷积网络结构的层数分别为 3,6,10 和 12。可以看出,设置为 6 层时效果最好,相比基准系统提高了约 0.96 个 BLEU 值;设置 3 层时可提高约 0.91 个 BLEU 值;设置为 10 层时可提高约 0.94 个 BLEU 值;当逐渐加深到 12 层时只能提高约 0.7 个 BLEU 值。根据表 1 的实验结果可知,当多层卷积网络设置为 6 层、10 层时,翻译质量最好,因此,我们固定卷积层数,设置卷积核为  $3 \times 1$  进行对比实验,实验结果如表 2 所列。其中,设置为 6 层时可以提高约 0.98 个 BLEU 值,超过了卷积核为  $2 \times 1$  时的实验效果,但是设置为 10 层时只能提高 0.84 个 BLEU 值,设置为 12 层时提高约 0.83 个 BLEU 值。

综合实验结果,可以得到以下结论:

1)与基准的神经机器翻译系统相比,融合了注意力卷积框架的神经机器翻译的效果有所改善。

2)实验中 2 元词组的特征相对于 3 元词组更加明显,采用大小为  $2 \times 1$  的卷积核时翻译效果优于采用大小为  $3 \times 1$  的卷积核。

3)建立多层卷积网络结构时,层数不能过高或者过低,设置为 6 层和 10 层时的翻译效果优于更低层或者更高层的结果,随着层数的加深,机器翻译的效果有下降的趋势。

### 5.2.2 运行速度

本节对模型运行速度进行考量。如表 3 所列,在同样的参数配置下,对基准系统和添加了大小为  $3 \times 1$  的卷积核的多层卷积系统的运算速度进行统计。

表 3 不同模型的运算速度

Table 3 Operation speed of different models

实验系统	运行速度/(词/秒)
基准系统	8000
+6 层网络结构	6400
+10 层网络结构	6100
+12 层网络结构	5500

由表 3 可以看出,添加了多层卷积网络后,系统整体的运行速度有所下降,层数越高则速度越慢。但是总体来看,由于卷积运算的并行性,多层卷积的运行速度仍在可控范围内。

### 5.2.3 注意力信息

本节对多层卷积网络的注意力信息进行对比。如图 5 所示,对英文语句都使用小写处理,采用卷积核配置为  $2 \times 1$  且卷积网络设置为 6 层的模型,使用 Draw<sup>1)</sup> 画出注意力分布。颜色越亮的地方表示注意力信息越强。

源端输入语句为“不过,美国股市仍然继续下滑。”,对应目标端语句为“however, the us stock market continues to slide.”,语句末尾新增的“<s>”表示语句结束的符号。图 5(a)给出了使用初始的神经机器翻译框架得到的注意力分布信息,图 5(b)~图 5(d)给出了经历多层卷积网络结构的注意力信息。可以看出,随着卷积层数的增加,高层的卷积网络形成了愈加明显的短语级别的注意力信息,扩大了信息范围,短语之间的关联性不断增强,从而构造出一个具有层级信息的注意力分布。

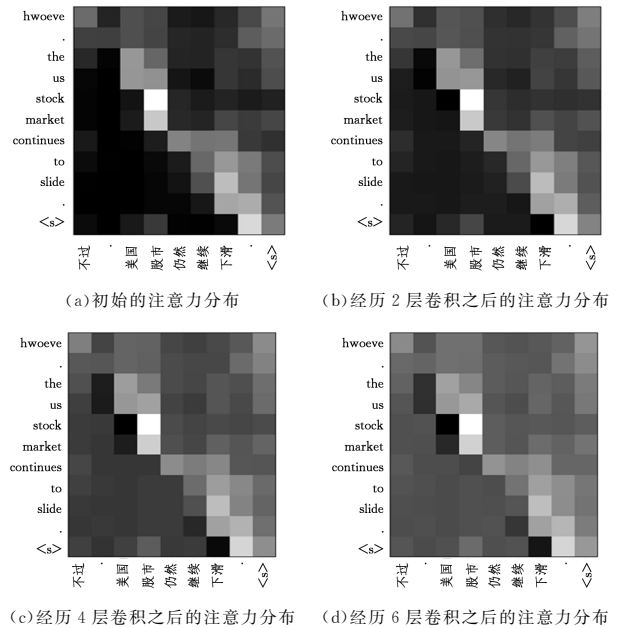


图 5 初始的注意力分布和经过多层卷积的注意力分布信息

Fig. 5 Initial attention distribution and attention distribution after different multi-layer convolution

**结束语** 本文针对现有的神经机器翻译结构提出了基于注意力卷积的神经机器翻译方法,通过对注意力机制建立多层卷积网络结构,获取语句中的短语结构信息,来增强语句前后的上下文依赖关系,将注意力信息从单词级别逐渐提升为短语级别。首先,建立多层卷积网络结构对每一时刻的注意力信息进行建模以提取特征;然后,通过门控机制对得到的信息进行限制和优化;最后,将多层卷积网络结构与基本的神经机器翻译结构相融合。实验结果显示,本文提出的方法可以有效地提高机器翻译的质量。

值得注意的是,本文使用的融合方法相对简单,没有将新得到的上下文信息和原有的上下文信息进行对比,提取出最优化的信息数据。并且,本文只对注意力信息多层卷积网络建模,未考虑到源端语句位置信息等内容的使用。在未来的工作中,我们将从以下几个方面进行深入研究:

1)如何更好地将新得到的上下文向量融入到原有的框架中,例如使用注意力机制将原有的和新得到的上下文向量信息融合;

2)在注意力层考虑语句信息和位置信息对注意力的影响,充分利用源端信息。

## 参考文献

- [1] FENG Z W. Studies of SCI-Tech Translation[M]. Beijing: China Translation Corporation, 2004. (in Chinese)  
冯志伟. 机器翻译研究[M]. 北京: 中国对外翻译出版公司, 2004.
- [2] 宗成庆. 统计自然语言处理[M]. 北京: 清华大学出版社, 2008.
- [3] LIU Q. Syntax-based Statistical Machine Translation Models and Approaches[J]. Journal of Chinese Information Processing, 2011, 25(6): 63-71. (in Chinese)

<sup>1)</sup> <https://github.com/liesun1994/Draw>

- 刘群. 基于句法的统计机器翻译模型与方法[J]. 中文信息学报, 2011, 25(6): 63-71.
- [4] KOEHN P, OCH F J, MARCU D. Statistical Phrase-based Translation[C]// Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1. Association for Computational Linguistics, 2003: 48-54.
- [5] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. arXiv: 1409.0473, 2014.
- [6] LUONG M T, PHAM H, MANNING C D. Effective Approaches to Attention-based Neural Machine Translation[J]. arXiv: 1508.04025, 2015.
- [7] SUTSKEVER I, VINYALS O, LE Q V. Sequence to Sequence Learning with Neural Networks[C]// Advances in Neural Information Processing Systems. 2014: 3104-3112.
- [8] LI Y C, XIONG D Y, ZHANG M. A survey of Neural Machine Translation[J/OL]. Chinese Journal of Computers. <http://cjic.ict.ac.cn/online/bfpub/lyc-20171229152034.pdf>. (in Chinese)  
李亚超, 熊德意, 张民. 神经机器翻译综述[J/OL]. 计算机学报. <http://cjic.ict.ac.cn/online/bfpub/lyc-20171229152034.pdf>.
- [9] GEHRING J, AULI M, GRANGIER D, et al. Convolutional Sequence to Sequence Learning[J]. arXiv preprint arXiv: 1705.03122, 2017.
- [10] XU K, BA J, KIROS R, et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[C]// International Conference on Machine Learning. 2015: 2048-2057.
- [11] CHENG Y, WU H, WU H, et al. Agreement-based joint training for bidirectional attention-based neural machine translation[J]. arXiv preprint arXiv: 1512.04650, 2015.
- [12] TU Z, LU Z, LIU Y, et al. Modeling Coverage for Neural Machine Translation[J]. arXiv preprint arXiv: 1601.04811, 2016.
- [13] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]// Advances in Neural Information Processing Systems. 2017: 5998-6008.
- [14] CHO K, MERRIENBOER B, GULECHRE C, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation[J]. arXiv preprint arXiv: 1406.1078, 2014.
- [15] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [17] NEUBIG G, DYER C, GOLDBERG Y, et al. DyNet: The Dynamic Neural Network Toolkit[J]. arXiv: 1701.03980.
- [18] NEUBIG G. Lamtram: A toolkit for language and translation modeling using neural networks[OL]. <http://www.github.com/neubig/lamtram>, 2015.
- [19] PAPINENI K, ROUKOS S, WARD T, et al. Bleu: a Method for Automatic Evaluation of Machine Translation[C]// Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2002: 311-318.
- (上接第 209 页)
- [10] JIA Y K. Research and Design on Data Extraction in Multiple Data Sources[D]. Harbin: Harbin Engineering University, 2013. (in Chinese)  
贾艳凯. 多源异构增量数据抽取方法研究与设计[D]. 哈尔滨: 哈尔滨工程大学, 2013.
- [11] YANG L. Design and Implementation of Real-time data extraction Mechanism in data warehousing [D]. Beijing: Beijing University of Posts and Telecommunications, 2007. (in Chinese)  
杨乐. 数据仓库中实时抽取机制的研究与实现[D]. 北京: 北京邮电大学, 2007.
- [12] TAN G W, WU T. Study on Method of Data Warehouse Real-time Data Updating Based on Mechanism of CDC [J]. Computer Science, 2015, 42(S1): 546-548. (in Chinese)  
谭光玮, 武彤. 基于 CDC 机制的数据仓库实时数据更新方法研究[J]. 计算机科学, 2015, 42(S1): 546-548.
- [13] ZOU X X, JIA W J, PAN J H. Research of Log-based Change Data Capture [J]. Journal of Chinese Computer Systems, 2012, 33(3): 531-536. (in Chinese)  
邹先霞, 贾维嘉, 潘久辉. 基于数据库日志的变化数据捕获研究[J]. 小型微型计算机系统, 2012, 33(3): 531-536.
- [14] JAIN T, SALUJA S. Refreshing Datawarehouse in Near Real-Time[J]. International Journal of Computer Applications, 2012, 46(18): 24-29.
- [15] WANG Y B, RAO X R, HE P. Incremental database synchronization update mechanism under heterogeneous environment [J]. Computer Engineering and Design, 2011, 32(3): 948-951. (in Chinese)  
王玉标, 饶锡如, 何盼. 异构环境下数据库增量同步更新机制[J]. 计算机工程与设计, 2011, 32(3): 948-951.
- [16] CUI Y W, ZHOU J H. Research on Data Integration Based on KETTLE [J]. Computer Technology and Development, 2015(4): 153-157. (in Chinese)  
崔有文, 周金海. 基于 KETTLE 的数据集成研究[J]. 计算机技术与发展, 2015(4): 153-157.
- [17] LIU X Q, WU G, DENG H P. Data deduplication in Web information integration[J]. Journal of Computer Applications, 2013, 33(9): 2493-2496. (in Chinese)  
刘雪琼, 武刚, 邓厚平. Web 信息整合中的数据去重方法[J]. 计算机应用, 2013, 33(9): 2493-2496.
- [18] TANK D M. Reducing ETL Load Times by a New Data Integration Approach for Real-time Business Intelligence[J]. International Journal of Engineering Innovations & Research, 2012, 1(2): 56-60.