

# 用 Recursive Neural Networks 得到分析树



不会停的蜗牛 (/u/7b67af2e61b3) + 关注

2016.09.23 12:31\* 字数 1246 阅读 1967 评论 1 喜欢 6

(/u/7b67af2e61b3)

CS224d-Day 10:

Recursive neural networks -- for parsing

课程链接 ([https://link.jianshu.com?](https://link.jianshu.com?t=https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.html)

[t=https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.html](https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.html))

视频链接 ([https://link.jianshu.com?t=https://www.youtube.com/watch?v=D4j\\_9Jn-E8g&index=10&list=PLIJy-eBtNFt4CSVWYqscHDdP58M3zFHIG](https://link.jianshu.com?t=https://www.youtube.com/watch?v=D4j_9Jn-E8g&index=10&list=PLIJy-eBtNFt4CSVWYqscHDdP58M3zFHIG))

课件链接 ([https://link.jianshu.com?](https://link.jianshu.com?t=https://web.archive.org/web/20160313082614/https://cs224d.stanford.edu/lectures/CS224d-Lecture9.pdf)

[t=https://web.archive.org/web/20160313082614/https://cs224d.stanford.edu/lectures/CS224d-Lecture9.pdf](https://web.archive.org/web/20160313082614/https://cs224d.stanford.edu/lectures/CS224d-Lecture9.pdf))

本文结构：

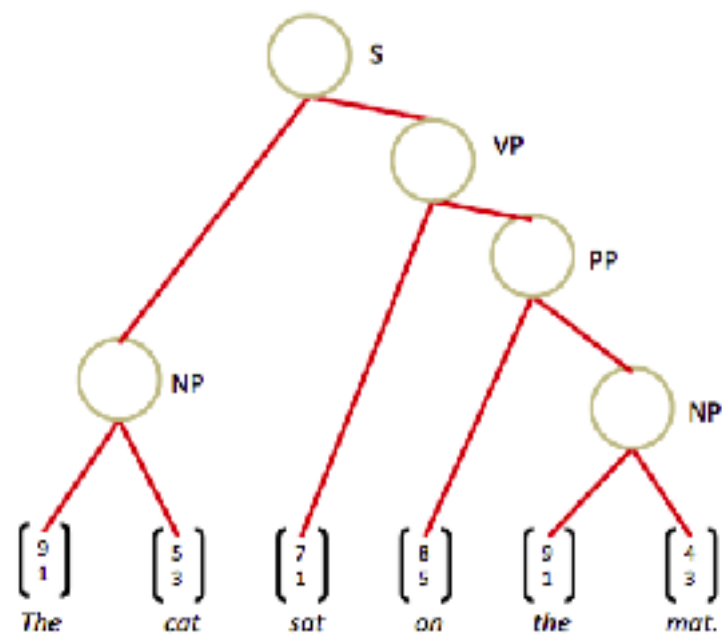
- Recursive NN 是什么
- Recursive Neural Networks 和 Recurrent Neural Networks
- Recursive NN 可以用来做什么
- 怎样做到的
- 算法代码



## • Recursive NN 是什么

Recursive Neural Networks 可以用来表达长句子，将一个句子映射到向量空间。

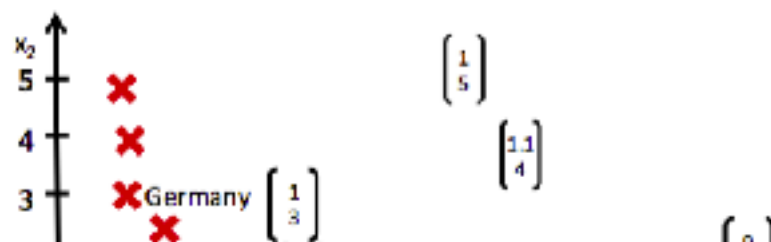
通过分析出句子的 parsing tree 的结构，把一个句子拆分成几个小组成单元，然后可以替换其中的一些部分，进而得到一些相似的句子，比如把这个 NP 名词短语，换成另一个 NP 名词短语。



这些句子由不同的短语组成，但是表达的意思却是一样的，在向量空间中，它们的距离也会很近。例如 'the country of my birth' 和 'the place where I was born' 意思一样，向量空间上的表达也就很近。

(/apps/download?  
utm\_source=sbc)



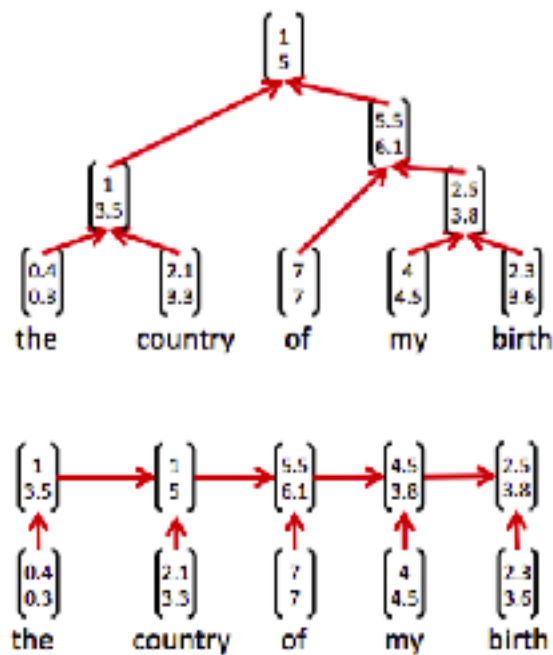


(/apps/download?  
utm\_source=sbc)

### • Recursive Neural Networks 和 Recurrent Neural Networks

想要分析数据的 hierarchical structure 的时候，Recursive NN 要比 Recurrent NN 更有效一些。

Recurrent NN 是 Recursive NN 的一种特殊形式，一个链可以写成一棵树的形式。



Recursive NN 可以返回树上每个节点的向量表达，Recurrent NN 在任何时候返回一句话的向量。

- **Recursive NN 可以用来做什么**

Recursive NN 可以用一个很好的方式来描述句子。

可以识别句子中的成分，可以通过替换组件来形成同样合理的句子，可以处理歧义问题，分析句子的语法结构，语义结构，理解一段话的指代词的含义。

可以学习到一个句子里哪个组成成分更重要，比如VP比NP更重要。可以学习到哪几个句子意思相近。

1. 当我们需要学习句子的结构的时候，会用 Recursive Neural Networks 来的到 parsing tree。
2. 也可以用来做 sentiment analysis，因为这个情感喜好的结果，不仅仅和单词本身有关，还和句子组成和顺序有关。
3. 还可以用来分析图片的组成，比如它可以分析出房顶，二层楼，一层楼，并组成一个房子。

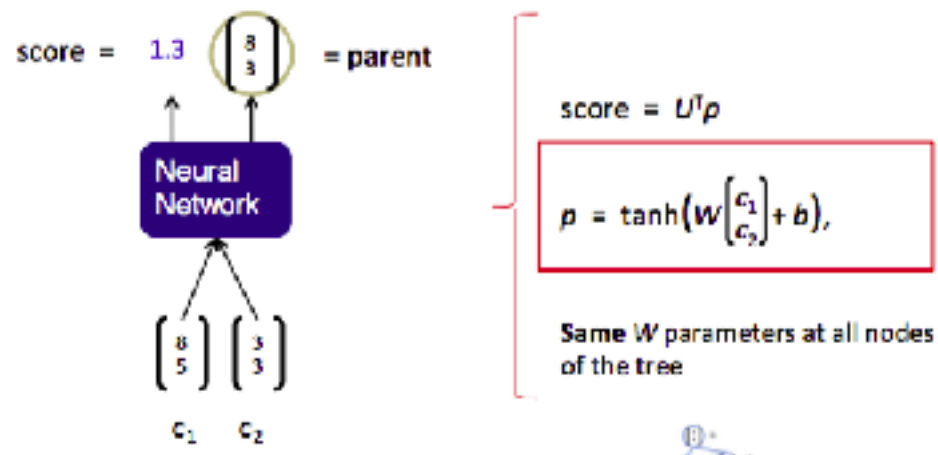
(/apps/download?  
utm\_source=sbc)



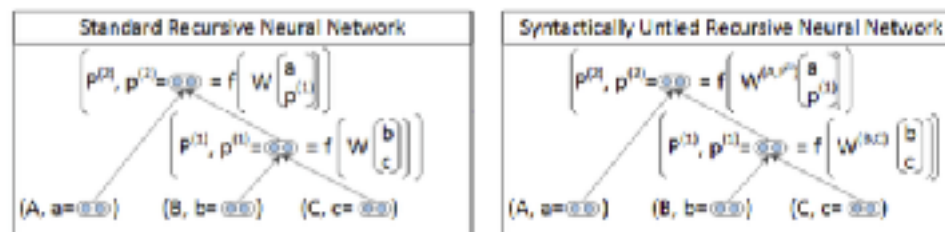
### • 怎样做到的

RNN 的输入是句子，输出是一个 parse 树结构。

下图是一个最标准的神经层，W 在整个网络中是一样的。



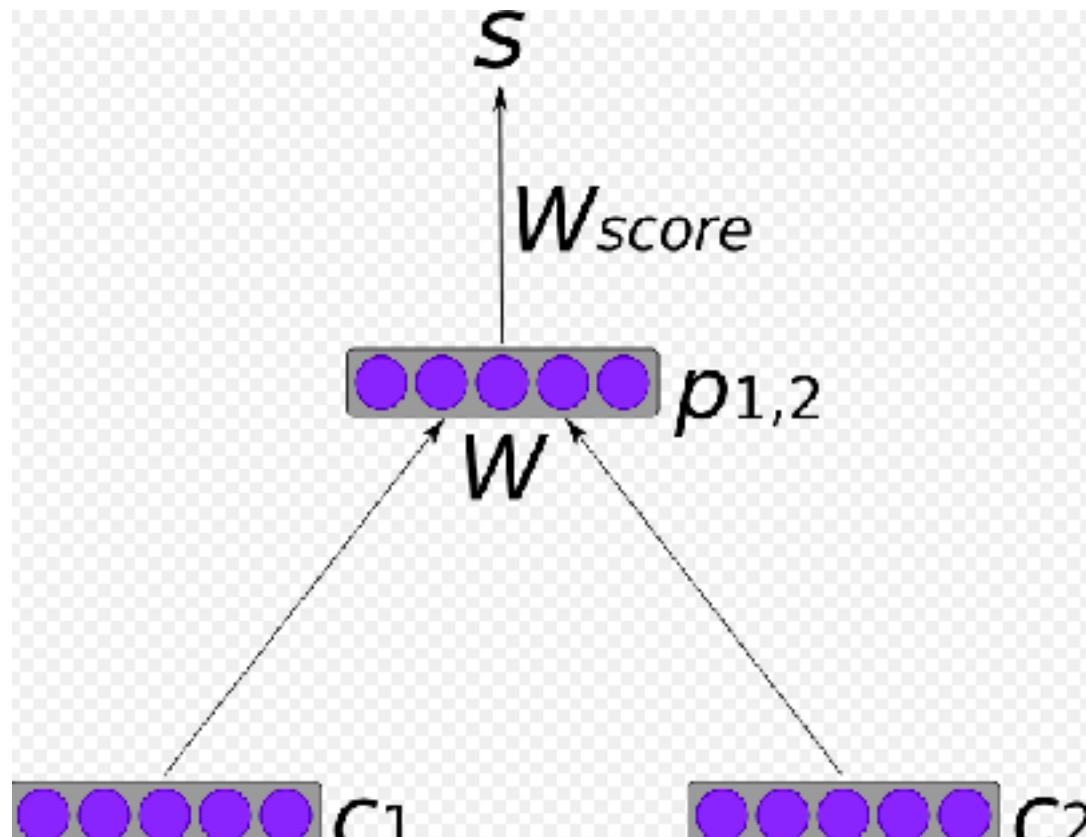
有个拓展模型 Syntactically-United RNN，是根据不同的组成成分使用不同的 W



RNN由三部分组成，根，左叶子，右叶子，也就是一个 Binary Tree。

它的每个节点上由一些神经元组成，神经元的个数由句子的复杂程度决定。

叶子是接收数据的，也就是向量，根是分类和评分的。



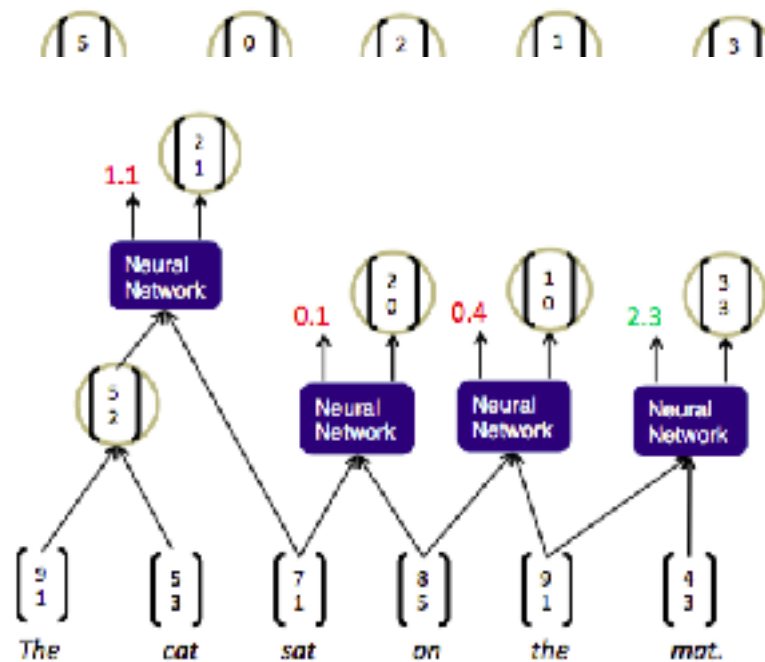
(/apps/download?  
utm\_source=sbc) ×

第一步，会先把句子的结构学习出来。

单词两两组合，进行评分，再作为一个整体，和后面的一个单词组合，再评分。

两个单词如果应该放在一起，就会得到高分，否则分数较低。





(/apps/download?utm\_source=sbc) ×

一个句子会得到多个结构，用 Greedy 选择其中分数最高的，作为最终的树结构。

用 Max Margin 来学习最优的树结构。每个  $i$  代表一个句子， $A(x_i)$  是包含  $x_i$  的所有可能的树，当  $y$  与  $y_i$  一样时， $\delta = 0$ 。



(/apps/download?  
utm\_source=sbc) ×

模型训练时，通过比较 labeled 数据，比较合适的结构和预测的结构，用 BTS 使误差达到最小。





第二步，再为结构中的每个小部分找到合适的语法标签，判断是什么成分的短语，是NP名词短语，VP动词短语，还是PP介词短语等。

- 算法代码

定义线性的神经元，做内积  $W(\text{left} + \text{right}) + b$

用 softmax 对每个点做 classify

`node.probs -= np.max(node.probs)` 这个技巧可以保证stable

(/apps/download?  
utm\_source=sbc) ×

下面的代码就是计算红色框里的式子



(/apps/download?  
utm\_source=src) ×

## [cs224d]

Day 1. 深度学习与自然语言处理 主要概念一览

(<https://www.jianshu.com/p/6993edef96e4>)

Day 2. TensorFlow 入门 (<https://www.jianshu.com/p/6766fbc43b9>)

Day 3. word2vec 模型思想和代码实现 (<https://www.jianshu.com/p/86134284fa14>)

Day 4. 怎样做情感分析 (<https://www.jianshu.com/p/1909031bb1f2>)

Day 5. CS224d - Day 5: RNN快速入门 (<https://www.jianshu.com/p/bf9ddfb21b07>)

Day 6. 一文学会用 Tensorflow 搭建神经网络

(<https://www.jianshu.com/p/e112012a4b2d>)

Day 7. 用深度神经网络处理NER命名实体识别问题

(<https://www.jianshu.com/p/581832f2c458>)

Day 8. 用 RNN 训练语言模型生成文本 (<https://www.jianshu.com/p/b4c5ff7c450f>)

Day 9. RNN与机器翻译 (<https://www.jianshu.com/p/23b46605857e>)

Day 10. 用 Recursive Neural Networks 得到分析树

(<https://www.jianshu.com/p/403665b55cd4>)

Day 11. RNN的高级应用 (<https://www.jianshu.com/p/0e840f92b532>)

---

我是 不会停的蜗牛 Alice

85后全职主妇

喜欢人工智能，行动派

创造力，思考力，学习力提升修炼进行中

欢迎您的喜欢，关注和评论！

📖 技术博文 (/nb/5173140)

举报文章 © 著作权归作者所有





不会停的蜗牛 (/u/7b67af2e61b3)

写了 224835 字，被 3243 人关注，获得了 1969 个喜欢

(/u/7b67af2e61b3)

+ 关注

我是 Alice 喜欢人工智能，行动派 创造力，思考力，学习力提升修炼进行中 欢迎志同道合的小伙伴们和我一...

(/apps/download?  
utm\_source=sbc)



喜欢就点赞，有用就随意打赏吧 😊 Run with AI !

赞赏支持

♡ 喜欢 (/sign\_in?utm\_source=desktop&utm\_medium=not-signed-in-like-button) | 6



更多分享

(http://cwb.assets.jianshu.io/notes/images/5949384/weibo/image\_5



下载简书 App ▶  
随时随地发现和创作内容



(/apps/download?utm\_source=nbc)



登录 (/sign\_in?utm\_source=desktop&utm\_medium=not-signed-in-comment-form)

1条评论

只看作者

按喜欢排序 按时间正序 按时间倒序



blue1sky (/u/25959dc6936b)

2楼 · 2017.07.24 16:07

(/u/25959dc6936b)



请问您有没有完整的代码，能够训练得到recursive网络的参数？

👍 赞    💬 回复

(/apps/download?  
utm\_source=sbc) ×

被以下专题收入，发现更多相似内容



数据科学家 (/c/0adc32d3cf07?utm\_source=desktop&utm\_medium=notes-included-collection)



坚持写作100天 (/c/6f43264f8299?utm\_source=desktop&utm\_medium=notes-included-collection)



程序员 (/c/NEt52a?utm\_source=desktop&utm\_medium=notes-included-collection)



NLP (/c/173cbecfcfc2?utm\_source=desktop&utm\_medium=notes-included-collection)

## Deep Learning in NLP-词向量和语言模型 (/p/d46609947489?utm\_campa...

Deep Learning 算法已经在图像和音频领域取得了惊人的成果，但是在 NLP 领域中尚未见到如此激动人心的结果。关于这个原因，引一条我比较赞同的微博。@王威廉：Steve Renals算了一下icassp录取文章题目中...



stonelin3935 (/u/6c767c086b0d?utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

## 今日学术视野(2017.11.16) (/p/d86b136b3f41?utm\_campaign=maleskine...

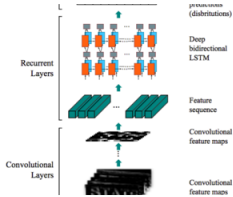
cs.AI - 人工智能cs.CL - 计算与语言cs.CR - 加密与安全cs.CV - 机器视觉与模式识别cs.CY - 计算与社会cs.DC - 分布式、并行与集群计算cs.DS - 数据结构与算法cs.HC - 人机接口cs.IR - 信息检索cs.IT - 信息论...



爱可可\_爱生活 (/u/ZQtGe6?utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)




(/p/9a3114e91c34?



(/apps/download?utm\_source=sbc)

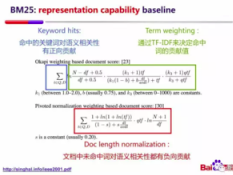
utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)  
**CRNN论文翻译——中英文对照 (/p/9a3114e91c34?utm\_campaign=males...**

文章作者：Tyan博客：noahsnail.com | CSDN | 简书 An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to ...

 SnailTyan (/u/7731e83f3a4e?


utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/3a9f49834c4a?



utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)  
**浅谈智能搜索和对话式OS (/p/3a9f49834c4a?utm\_campaign=maleskine&...**


前面的文章主要从理论的角度介绍了自然语言人机对话系统所可能涉及到的多个领域的经典模型和基础知识。这篇文章，甚至之后的文章，会从更贴近业务的角度来写，侧重于介绍一些与自然语言问答业务密切相..

 我偏笑\_NSNirvana (/u/2293f85dc197?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

**今日学术视野(2016.11.23) (/p/c2cbf74ec038?utm\_campaign=maleskine&..**


cs.AI - 人工智能cs.CL - 计算与语言cs.CR - 加密与安全cs.CV - 机器视觉与模式识别cs.DB - 数据库cs.DC - 分布式、并行与集群计算cs.DS - 数据结构与算法cs.HC - 人机接口cs.IR - 信息检索cs.LG - 自动学习c...

 爱可可\_爱生活 (/u/ZQtGe6?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

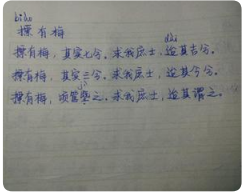
**模特被大火毁容，男友依旧视她为唯一的梦中情人 (/p/ada934b67613?utm\_...**

澳洲一位女子是一名健身爱好者，因为身材姣好，也曾做过模特。2011年，她参加马拉松的途中遭遇山火，全身被烧伤65%。原本靓丽的外表被烧得面目全非，她还失去了7根手指。但她相处多年的男朋友一直不离...

 破镜重圆公司 (/u/425e2a68b97d?)

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation) (/apps/download?utm\_source=sbc) ×

(/p/f1bffc858cea?



utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

**打卡第二十四天 (/p/f1bffc858cea?utm\_campaign=maleskine&utm\_conte...**

今天做了很不好的事情，以后不能、也不会，总之就是不可以。常常以为自己很是需要拓宽视野，活得太闭塞。听课两个小时左右吧。大概是别人一天上课的时间，我听了四五六七八天，还没吃透的那种，智商...

 飘飘西 (/u/ea953f6865ea?)

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)


(/p/bf1f2345b097?



utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

**《七绝·垂钓》 (/p/bf1f2345b097?utm\_campaign=maleskine&utm\_conte...**

碧水悠悠柳色新，也傍钓客在湖滨；悬钩半日无鱼影，只闻鸣蝉乱我神。


 清净菩提 (/u/d23c44565e3b?)

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

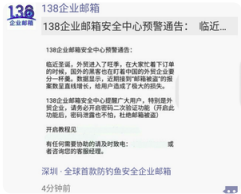
**写给儿子的信——来自突然的灵感 (/p/4566ed97a7c6?utm\_campaign=mal...**

2017年9月22日 儿子，在睡梦中听到你的咳嗽，妈妈的手抚摸着你的背，在心里默默的对你说：孩子，你是健康的，你是被爱的，你是安全的，你是强壮的，你是开心，你是快乐，你是喜悦，你是丰盛，你是爱，你..



 内在的宁静 (/u/f28a23f4e866?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)


(/p/a76d02feea3b?



(/apps/download?  
utm\_source=sbc) ×

utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)  
**138企业邮箱安全中心预警通告 (/p/a76d02feea3b?utm\_campaign=malesk...**

138企业邮箱安全中心预警通告：临近圣诞，外贸进入了旺季，在大家忙着下订单的时候，国外的黑客也在盯着中国的外贸企业要分一杯羹。数据显示，近期接到“邮箱被盗”的报案数呈直线增长，给用户造成了极大...

 网谈世界 (/u/84f106db4dd1?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

