首先，我是在NBLucene的论文的未来的工作去找这篇论文的，NBLucene上面说希望能够实现在以GPU为系统架构的搜索引擎的实现。

而自己找的这篇论文正是 以是首次以GPU系统架构搜索引擎的讨论，可以说是对第一篇论文的延伸内容。

两篇论文的联系：

大致都介绍了搜索引擎的相关步骤，比如说建立倒排索引，索引压缩的方法，以及以并行方式解码的相关规则，

不同：

1：NBLucene侧重点在于跟已有的方法进行改良，然后看提升其效率，比如说将语言改为C/C++编写，由于C/C++直接控制内存的分配与释放，能够直接与处于底层的硬件打交道，效率显然可以提高很多。而这篇论文，更多的是开启了一个新的篇章，因为之前没有关于GPU架构的搜索引擎的实现，所以也没有效率提升的一部分，更多的是以在CPU为系统架构的比较工作。

2：关于列表合并检索和分离检索，GPU的这篇论文做了大量的介绍。可以说GPU的这篇文章更加详细的介绍了搜索引擎内部是怎么具体实现完成搜索功能的。

3：NBLucene详细介绍了以cache的逻辑实现，而在以GPU系统架构中并没有提相关内

容，可能是因为在设计GPU时目的就是为了让其做大量运算，而不是以控制工作为目的。

对于未来的工作我的想法如下：

1：近期目标：

由于我是初步入门搜索引擎的，经过了两个星期的理论知识，正好我想通过这次的理论先锻炼一下实际编码C/C++的能力。大四或者毕设方向希望完成一个能够查询论文的搜索引擎。这样我觉得也对大学四年学的计算机基础知识有个总结。（本来我毕设是想做Android方向的即时通讯模块的，但正如刘老师在组会上所说的，我们现在学的一些东西比如说Android ，或者Java的相关框架（Spring+Struts+Hibernate）的知识，，对于我们了解计算机内部一点帮助都没有,这些内容只是在给个饭碗，让我们出去能够找到工作仅此而已，所以果断放弃）

2：在研究生阶段：

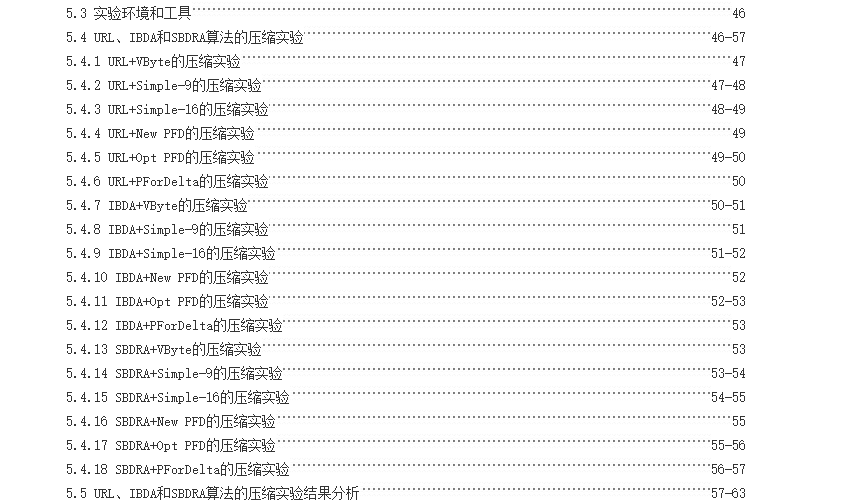
蹭课（我也知道我们软件工程阶段本科期间学习的内容比较水，比如说我们连编译原理这门课都没开。。。所以想去蹭一下南开本科生计算机的相关课程） - > 如编译原理 和 计算机组成原理 需要对计算机内部有着清晰的认识

3: 论文方向：

1：可以对以GPU为基础的索引分层，和 提前终止的相关方法需要进一步的研究。

2: 我在查资料的同时，看到了索引校正问题，这两篇论文都没有讨论相应的情况，比如说错别字改正等有关查询的功能，以后也可以在这个方向去完善。

3：可以讨论一下这两篇论文压缩算法的效率，如我自己找的论文运用了两种压缩算法，比如Rice Coding 和 pForDelta方法，而Nblucene这篇论文介绍了NewPFor，可以讨论一下各个压缩算法在不同情况下的适用范围。如这篇论文：



4：对机器学习和数据挖掘的相关步骤，需要更深一步的学习。

5：这里也可以对继续NBLucene的提的脚步，以C/C++为主要语言，完成在GPU中系统架构的实现。