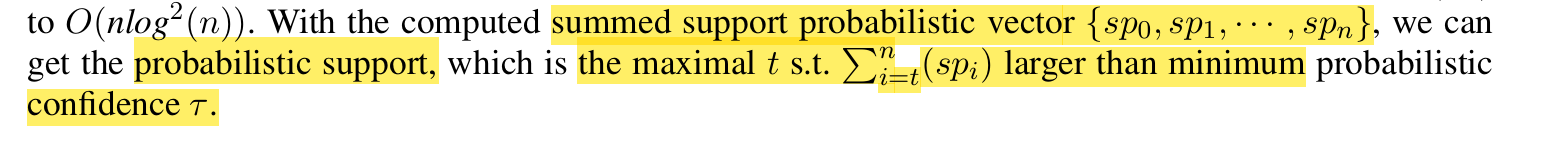
**Probabilistic maximal frequent itemset mining methods over uncertain databases**

**4.2. Probabilistic support computing:**

Because the summed support probabilistic vector of itemset x in two transactions T1 and T2 is the convolution of that in T1 and that in T2, the divide-and-conquer method proposed in [20] is also employed in our paper.

That is, the uncertain database will be split into two parts to separately compute the summed support probabilistic vector, and this operation will be recursively conducted until the sub database has only one transaction. The convolution can be computed with a Fast Fourier Transform, which, given the size of the uncertain database n, will efficiently reduce the time complexity from O(n2) to O(nlog2(n)).

With the computed summed support probabilistic vector , we can get the probabilistic support, which is maximal of t so that: .



**Efficient weighted probabilistic frequent itemset mining in uncertain databases**

Result of this paper is collection of weighted probabilistic frequent itemset:

In paper, people propose a breath-first mining algorithm, so I store result in list:

[[size-1 frequent itemset],

[size-2 frequent itemset],

…

[size-n frequent itemset]]

**(Probabilistic Maximal Frequent Itemset):** Given the uncertain database, minimum support, minimum confidence, an itemset is probabilistic maximal frequent itemset if it is probabilistic frequent itemset and it isn’t covered by another probabilistic frequent itemset, denoted: .

**Theorem 3: (Anti-monotonicity property for weighted PFI)** If itemset **x** is w-PFI with size , then there is at least one itemset is w-PFI.

**How to find weighted probabilistic maximal frequent itemset in this?**

for (0 -> n-2) {

size-k itemset list,

size-(k+1) itemset list

if (not found any element in size-(k+1) covered an element in size-k) {

size-k element is maximal;

}

}

size-n frequent itemset is definitely maximal.