**iTunes Crawler**

* Get the iTunes U Universities and colleges’ link.
* Retrieve the XML file of URL by setting Use-Agent to iTunes.
* Parse the XML file to obtain main URL of universities and save in table MAIN\_TABLE.

**Algorithm:**

crawler

{

Loop1:

Remove first URL from MAIN\_TABLE

do parsing

/\*\*

\*Span threads to carry out Loop2

\*BFS is performed

\*/

Loop2:

Remove first URL from WAITING\_TABLE //Atomic operation

do parsing

Place URL in CRAWLED\_TABLE

Put the main URL in CRAWLED\_TABLE

}

parsing

{

Get the XML document of URL and save the file

Check if it contains TrackList

if(present)

{

save all tracks with their details in the TRACKS table

}

Retrieve URLs in file and place it WAITING\_TABLE

}

**Tables used:**

MAIN\_TABLE - Contains main URL of Universities to be crawled

WAITING\_TABLE - Contains list of sites in an University to be crawled

CRAWLED\_TABLE - Contains list of sites in an University which have already been crawled

TRACKS - contains the list of saved tracks from the crawled university