

Re: Dealing with Web Mining on Apache Log

Source:

<http://www.tech-archive.net/Archive/SQL-Server/microsoft.public.sqlserver.datamining/2004-03/0054.html>

From: Peter Kim [MS] (peterkim_at_online.microsoft.com)

Date: 03/10/04

Date: Tue, 9 Mar 2004 17:13:42 -0800

Hmm... I thought that you could do the loading with DTS and SQL server without using Commerce Server. It's just that the MSDN article was created for Commerce Server customers. I might be wrong, though.

--

Peter Kim

This posting is provided "AS IS" with no warranties, and confers no rights.

<anonymous@discussions.microsoft.com> wrote in message

news:985b01c405d9\$2329f190\$a501280a@phx.gbl...

> Dear Peter,

>

> Many Thank!

> I've tried the trail version of the Commerce Server,

> however, the function for loading log is disabled...so

> disappointed..

>

> Cheers,

> Jenny

> >-----Original Message-----

> >Actually, I found an MSDN entry showing how to use SQL

> >DTS to load web log

> >files to your SQL Server data warehouse. Hope you find

> >it useful:

> >

> >[http://msdn.microsoft.com/library/default.asp?](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/csvr2002/htm/cs_mmc_2datawarehouse_dayx.asp)

> >url=/library/en-

> >us/csvr2002/htm/cs_mmc_2datawarehouse_dayx.asp

> >

> >--

> >Peter Kim

> >This posting is provided "AS IS" with no warranties, and

> >confers no rights.

> >

> >"Peter Kim [MS]" <peterkim@online.microsoft.com> wrote

> >in message

> >news:OCWCxm\$\$DHA.2348@TK2MSFTNGP09.phx.gbl...

> >> The first question is data preparation issue that the

> >>MS Analysis Services

> >>DM component doesn't directly addresses. I'm

> >>forwarding to DTS group in

> >>case

> >>they have a suggestion.

> >>

> >>Once you load, the first two questions could be

> >>answered by simple SQL

microsoft.public.sqlserver.datamining: Re: Dealing with Web Mining on Apache Log

```
> >> queries. The other two are more sequence analysis
> problems. MS Analysis
> >> Services 2000 DM component doesn't directly support
> sequence analysis, but
> >> you could use decision trees and clustering algorithms
> to analyze the log
> >> without ordering being modelled. I believe DBMiner
> (www.dbminer.com) has an
> >> implementation for sequence analysis as an aggregated
> provider of MS
> >> Analysis Services.
> >>
> >> --
> >> Peter Kim
> >> This posting is provided "AS IS" with no warranties,
> and confers no
> >rights.
> >>
> >> "jenny" <anonymous@discussions.microsoft.com> wrote in
> message
> >> news:48cc01c3fedb$91cd2720$a101280a@phx.gbl...
> >> > I am new to Data Mining but I need to develop a tool
> for
> >> > mining the apache log file for my final year project.
> >> > I have several questions to ask.
> >> >
> >> > 1. How can I insert apache log file like access_log
> and
> >> > referer_log into the sql server 2000, any tools can
> load
> >> > all
> >> > the data immediately?
> >> >
> >> > Format of the access_log:
> >> > 218.102.21.133 - - [01/Sep/2003:00:00:03
> >> > +0800] "GET /cslab/pics/d_hours.gif HTTP/1.1" 304 -
> >>
> > > "http://www.cs.cityu.edu.hk/cslab/left.html" "Mozilla/4
> .0
> >> > (compatible; MSIE 6.0; Windows NT 5.0)"
> >> >
> >> > Format of the referer_log:
> >> > http://www.cs.cityu.edu.hk/~fypms/student\_menu.cgi?
> >> > show_proposal.cgi -> /~fypms/student.cgi
> >> >
> >> > 2. Is it possible to get statistics result like
> >> > a. the greatest number of hits
> >> > b. where the potential applicants came from
> >> > c. where they will go after visiting the main page
> >> > d. show the users' access pattern over a certain
> period
> >> > time (equals to revisit the same pages)
> >> >
> >> > 3. Can ASP handle the above work with sql server
> 2000?
> >> >
> >> > Would anyone give me a hand on it?
> >>
> >>
> >
> >
> >.
```

microsoft.public.sqlserver.datamining: Re: Dealing with Web Mining on Apache Log

> >