Why Web Click Stream Analytic

Web click stream analytic provides critical insight into customer behavior in a business web site . Visitante provides a set of analytic solutions using web server logs.

Introduction

*Visitante* offers a suite of click stream analytic solutions on Hadoop and Storm. The solution is based on web server logs and does not require embedded java script in web pages.

Session Summarizer

First, log data is grouped by sessionID. Then session statistics are calculated, including number of pages visited, session duration.

It also detects if a flow was entered, whether or not the flow was completed. This feature is useful for identifying abandoned shopping carts.

It’s also possible to calculate statistics on a per user basis. Some the quantities calculated are average session duration, average number of pages per session, average gap between sessions.

Real time Analytic

Storm based real time analytic is also part of visitante. These analytic solutions are based on probabilistic aggregation algorithms. The solutions available are counting unique visitors and bounce rate.

Configuration

*Visitante* believes in giving power in the hands of the end users. There are many configuration parameters to control the behavior of the different algorithms

Data Format

The input data format is W3C compliant web server logs and the output data format is CSV.