

Real-Time Advertisement Targeting and Embedding in Web Sites

Rohit.Vijay.Vaidya
B.E (Information Technology) Lecturer
Dept. of. Information Technology
K.K.Wagh.I.E.E.R Nashik.
rohitev@gmail.com

Farahnaj.Abdulla. Inamdar
B.E.I.T
Dept. of. Information Technology
K.K.Wagh.I.E.E.R Nashik.
farahinamdar@gmail.com

Abstract

Online Advertisement generates a lot of revenue which funds websites online. The Search Engines and the advertisement publishers benefit from this business. The carriers of the Internet never get the share of this money though they carry the information. In this paper we are going to present a technology involved that will bring this revenue to the Internet Service Providers by introducing a layer of software that runs at the ISP.

We also would focus on the architecture of software system used in materializing this product. The software installed in the network of the Internet service provider will capture the requests made by the user as he surfs the internet. The pages will be delivered in a normal fashion, but before they are parsed by the software for keywords so that a profile of each users likes and interests is constructed. These profiles can then be helpful to target the advertisements accurately.

Keywords: RAT(Real Time Advertisement Targeting), Behavioral Targeting Optimizers, Online Advertisements, DNS (Domain Naming System) , Data Incinerator.

1. Introduction

RAT(Real Time Advertisement Targeting) is the technology that will reside in the Internet Service Providers network. It is a layer of software that resides on one of the ISP gateway. It has various real-time capabilities like user profiling and advertisement injection in near real time so that user browsing experience is consistent.

The existing Online Advertisement Industry mostly involves search engine giants because they are the ones who possess huge information(web pages) that can be parsed so that advertisements can be targeted accordingly. Though this has a advantage of targeting a contextual advertisement it has various limitations. The advertisers provide a pool of advertisements which are targeted by search engines using various targeting methods. Simplest of all is the Conflation Algorithm[1]. This simply are contextual advertisements and user behavior is not considered. Behavior implies surfing habits.

Is the user a random surfer? Or Is he interested in a particular entity? Which in turn might be something a advertiser wants to sell. With contextual behavior monitoring is not possible. The Internet Service Providers have an implicit advantage of being able to monitor the users traffic subject to various conditions.

With RAT you don't have to store anything as decisions are taken in real time with which advertisement are to be targeted. RAT will have to have to deal with the ISP which gives access to huge information that flows through its network. This switches the advertising from **page based to user based model**. Google Ad-sense works with the contextual ads scheme where the targeting rules are driven by page content and not User Profile. Plus Google has to store this information in its huge database for almost 2 years.

2. RAT System Overview

The RAT is as shown in the figure 1. As mentioned previously it is a layer of software that runs on one of the ISP Gateways.[4] It has namely Advertising Agencies, Behavioral Targeting Optimizers, Publishers, and the ISP. All these external entities to RAT have a web based access to RAT. The Advertising Agencies are the ones who provide the ISP with the pool of advertisements in either text format, flash or any image format. The ISP collaboration with the Publishers which are the web sites interested in the RAT scheme.

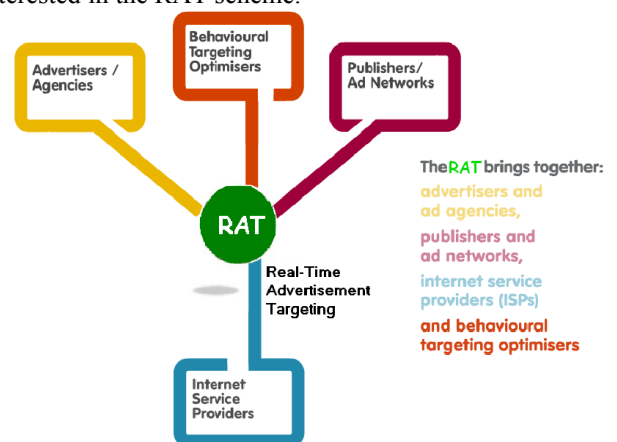


Figure1 :RAT Overview

Only websites with special headers which is authenticated by the RAT are available for run-time advertisement targeting. Rest are delivered as it is. The Behavioral Targeting Optimizers are the ones which provides rules to the RAT system to understand behavioral patterns. In simple terms it is a tweaking facility that allows to change certain RAT targeting functions.

2.1 How RAT Works

RAT assigns a random number to every user which becomes its identity while surfing the internet. The ISP being the carrier has access to huge pool of information flowing through its network. By tapping this network RAT is aware of the surfing habits of the user. It has a Instant Data Incinerator which immediately destroys any trace of users web page making the user anonymous. Prior to this it just parses the web page for certain index terms which can be used for user profiling. There is a Channel which categorizes the advertisements in the advertisement pool. This categorization proves to be helpful when targeting a advertisement.



Figure 2 Data Incinerator Preserving user anonymity.

2.2 How DNS Response are used ?

As the user surfs the internet the system generates various packets. These network packets are sniffed by RAT. The RAT system is always for the look for DNS [2](Domain naming System) Response packet which marks the beginning of any formal communication between a web server and the client (browser) . Using this information the RAT becomes aware of the two hosts involved in communication.[3] One which is the Web Server and the other the Client. This is where the profiling can begin for each individual user which includes the number of web

sites surfed. Surfing habits. Special inclination towards a keyword (topic) etc.

Following is the Ethernet II packet that runs on the local network. The structure is the same, but varies in size, regardless of whether it's a DNS Query or Response:

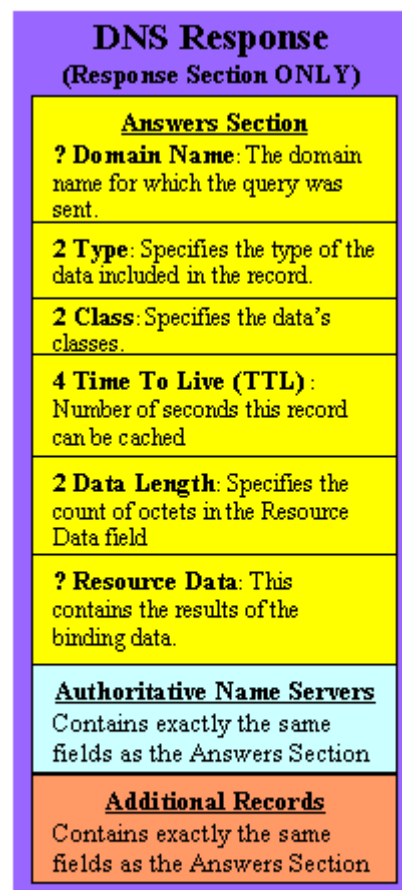
DNS Packet-Ethernet II Frame



A Simple Ethernet II Frame which is generated by a DNS Query or Response

Figure 3: TCP/IP Protocol Sequence

All the fields in the figure 4 are self explanatory.



All lengths shown are in bytes
 "?" indicates variable length field

Figure 4 DNS Response Packet

2.3 How are Advertisements Injected?

It is necessary for the web servers (publishers) to collaborate with the RAT(ISP) so that advertisements can be injected in a negotiated way. The RAT affiliated web servers are provided with special headers for identification and which intend to provide slots for advertisement injection.

The RAT system is a software running on the Gateway of the ISP. The Architecture of this software will involve multi-threading as real-time behavior is desired. The following steps are involved:

- 1) Create a Sniffer Thread
- 2) Create a Injector Thread
- 3) Sniff Packets from the network.
- 4) Send inputs to the decision module.
- 5) Send output of Decision Module to Injector
- 6) Inject Packets.
- 7) Goto 3

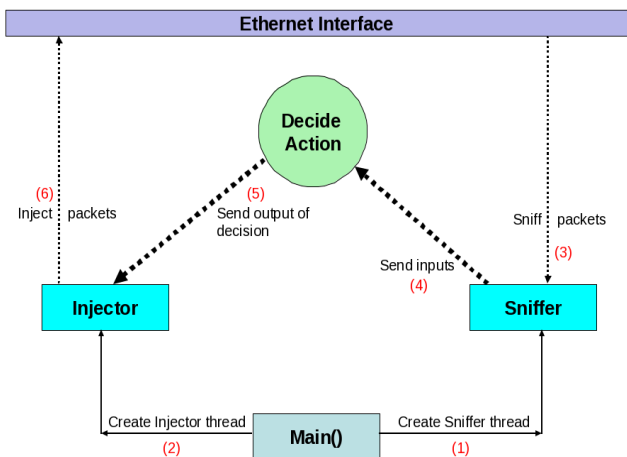


Figure 3 Sniffer Injector Logic

The RAT functions on the above mentioned logic. Its Decision module has inputs from user profiles. It then decides which advertisement to inject and when to do it. Using the above methodology every Advertisement is personalized. Users don't get any irrelevant ads. No browsing history or personally identifiable information is stored. User regain control over his privacy. Having said this we mean that RAT is a session oriented approach. From session to session user profiling would be done. Hence RAT is much more dynamic in nature and does not stick to one particular image of user. Human behavior by itself is dynamic and unpredictable.

3. Advantages Over Search Engines.

- 1) No Personal Data needs to be stored because of real-time aggregation of information.
- 2) Option for User-Opt is provided which is difficult process with Search Engines.
- 3) Browsing History is not stored whereas in Search Engines it may be stored for 12-24 months.
- 4) Accidental or malicious disclosure of stored information is possible with Search Engines. Not at all possible with RAT.
- 5) Ability to undo previous participation.

4. Future Enhancement

1. RAT can be tested with massive parallel computers to handle huge traffic in order to maintain QoS.
2. RAT can be enhanced to incorporate data mining so that patterns in the searches can be found.

5. Conclusion

Using RAT we can bring the online Advertisement business to the ISP which are the true carriers of information and have an implicit advantage of being able to intercept every traffic subject to certain terms and conditions. With RAT the user experience w.r.t to relevance in advertisements can be attained.

6. References

- [1] Modern IR Ricardo Baeza Yates
- [2] Data Communications and Networking Behroz A Forouzan
- [3] <http://www.firewall.cx/dns-query-format.php>
- [4] http://compnetworking.about.com/cs/isps/gr/aapr-cisco_ispe.htm