

Dr. sc. MISLAV ŠIMUNIĆ, viši asistent
Fakultet za turistički i hotelski menadžment u Opatiji, Sveučilište u Rijeci

UPRAVLJANJE POSLOVNIM REZULTATOM PROGRAMSKI UPRAVLJIVIM WEB ČVOROM, TEMELJENOM NA ANALIZI LOG DATOTEKE

Marketing putem Web stranica omogućuje jedinstvenu mogućnost stalnog prikupljanja podataka o ponašanju posjetitelja tih web stranica. Naime, za svaku je tvrtku, odnosno poslovni sustav od izuzetne važnosti prikupiti što više korisnih informacija i detalja o korisnicima svojih proizvoda i usluga. Činjenica je da se trendovi u svijetu sve dinamičnije mijenjaju i da ih je sve teže pratiti. U skladu sa time, jednom kreirane web stranice ne predstavljaju završen zadatak. Njihovo dotjerivanje, reorganiziranje mora biti stalan proces i to po mogućnostima u skladu sa dinamikom promjena na tržištu potražnje. No, to često i nije tako trivijalno izvedivo. Kroz daljnji tijek ovoga rada predstaviti će se model permanentnog sustavnog praćenja, analiziranja i reorganiziranja web čvora, «upravljivo» prema ponašanju, odnosno aktivnostima korisnika web stranica. Da bi se to postiglo potrebno je imati jasan uvidu u sve elemente web stranica i njihove karakteristike. Kako bi se moglo pratiti, analizirati i manipulirati sa web stranicama potrebno je jasno razumjeti **log datoteku**, njenu strukturu, stohastičnost, te iznaći načine za njeno logičko i funkcionalno programski vođeno uređivanje i iskorištavanje o čemu će više riječi biti u natrednim djelovima ovoga rada, nakon što se definira kvalitativni pristup izradu web čvora kao MmIM-a čiji oblik egzistiranja mora utjecati na poboljšanje poslovnog rezultata.

Ključne riječi: web čvor, log datoteka, analiza, dinamika, WebLogExplorer

1. IZRADA WEB ČVORA KAO MODELA (MMIM-A) I NJEGOVIH PODMODULA NA PRIMJERU HOTELA KAO POSLOVNOG SUSTAVA

Hotel kao poslovni sustav svoje promotivne aktivnosti uglavnom vrši putem mass medija, u sklopu kojih u posljednjih 10-ak godina Internet sve više dobiva na važnosti zbog mnogih prednosti, a koje se očituju u hipermedijskoj prirodi web stranica, mogućnosti interakcije korisnika sa poslovnim sustavom putem web stranica, te 24 satnoj dostupnosti korisnicima širom svijeta. Problem standardizacije uopće pa i MmIM-a za prezentaciju hotela na Internetu problem je koji će se trebati globalno riješiti u svrhu cjelokupnog

napretka ne samo Interneta kao sustava, već i svih onih kojima je on postao dio svakodnevnog poslovanja. Sama, naime problematika izrade MmIM-a prezentacije, sastoji se u složenosti definiranja, kako pojedinih entiteta modela tako i njihovih atributa odnosno varijabli koje ih određuju. Isto tako, problem predstavlja i određivanje pojedinih elemenata multimedije unutar modela, odnosno pojedinih entiteta, kojima se opisuju pojedine varijable kao konstantne veličine¹. U nastavku ovoga rada pojedini se entiteti, njihovi atributi i njihove varijable, različitom kombinacijom elemenata multimedije opisuju u detalje na primjeru hotela kao poslovnog sustava². Budući je osnovna djelatnost kako poslovnih sustava srodnih djelatnosti, pa tako i svih svjetskih hotela bez obzira na parametre koji ih determiniraju vrlo slična³, s velikom se sigurnošću mogu definirati entiteti i njihovi atributi kao elementi MmIM-a za prezentaciju hotela na Internetu i to⁴:

- [☰] Index Entitet. POČETNA STRANICA + SADRŽAJ.
- [①] Entitet 1. OPĆI PODACI/INFORMACIJE
- [📍] Entitet 2. LOKACIJA
- [↔] Entitet 3. SMJEŠTAJ
- [🍽️] Entitet 4. HRANA I PIĆE
- [🏖️] Entitet 5. REKREACIJA
- [🔗] Entitet 6. OSTALE USLUGE
- [↑] Entitet 7. OSOBLJE HOTELA
- [📄] Entitet 8. OBRAZAC ZA REZERVACIJE
- [🗨️] Entitet 9. OBRAZAC ZA REAKCIJE-INTERAKCIJU
- [🌐] Entitet 10. ZANIMLJIVI LINKOVI U DESTINACIJI I SVIJETU

U nastavku se ovoga poglavlja pojedinačno definiraju atributi odnosno varijable koje ih determiniraju.

¹ M. Randić, M. Šimunić: "Prezentacija hotela na Internetu", Hotelijerski fakultet Opatija, Hotelska kuća '98., str. 261-272, 1998.

² Svi se entiteti, varijable i atributi u disertaciji opisuju na hrvatskom jeziku. Ista bi sadržajno organizacijska koncepcija web čvora trebala egzistirati na svim toj destinaciji interesantnim jezicima (npr. engl., njem., tal., slov....)




³ Sung-Eon Kim, Thomas Shaw, Helmut Schneider: "Web site design benchmarking within industry groups" Internet Research, Bradford, 2003. Vol. 13, Iss. 1; str. 17

⁴ Šimunić, M.: "Suvremena komunikacija s tržištem – modeli prezentacije hotela na Internetu", Fakultet za turistički i hotelski menadžment, Opatija, 2000. – magistarski rad

Index Entitet: Početna stranica + sadržaj

Index Entitet predstavlja početnu stranicu kojom najčešće korisnik pristupa MmIM-u odnosno Web čvor-u određenog hotela (no to nije nužno budući se preraživanjem ulazi i u hijerarhijski niže razine), po čemu ovaj entitet dobiva karakteristike dominantnog entiteta. Početna stranica mora odasati prepoznatljivošću i jasnoćom tj. mora se odmah prepoznati da se radi o Web stranici određenoga hotela. Izgled početne stranice vrlo je važan iz razloga što ona najvećim dijelom trenutačno ostavlja pozitivan ili negativan utjecaj na korisnika, budući se upravo preko nje najčešće stječe prvi dojam o hotelu. Prepoznatljivost se može postići npr. poznatim logotipom ili nekom drugom slikom ili određenim prepoznatljivim znakom. No, glavni atributi, varijable i elementi multimedije kojima bi ovaj entitet trebao biti opisan prikazani su u slijedećoj tablici:

Tablica 1. 1: Opis Index Entiteta - početna stranica (home page)

Index entitet	atributi	varijable	elementi multimedije
POČETNA STRANICA + SADRŽAJ 	meni (sadržaj)	<ul style="list-style-type: none"> popis svih entiteta (uz mogućnost proširenja broja entiteta) 	 tekst  slika
	informacije za kontakt	<ul style="list-style-type: none"> puni naziv hotela, adresa, tel., fax., aktivni e-mail, ime i prezime osobe za kontakt 	
	"hotelski šarm"-prepoznatljivost	<ul style="list-style-type: none"> logotip hotela, puni naziv hotela kategorija hotela slika hotela 	

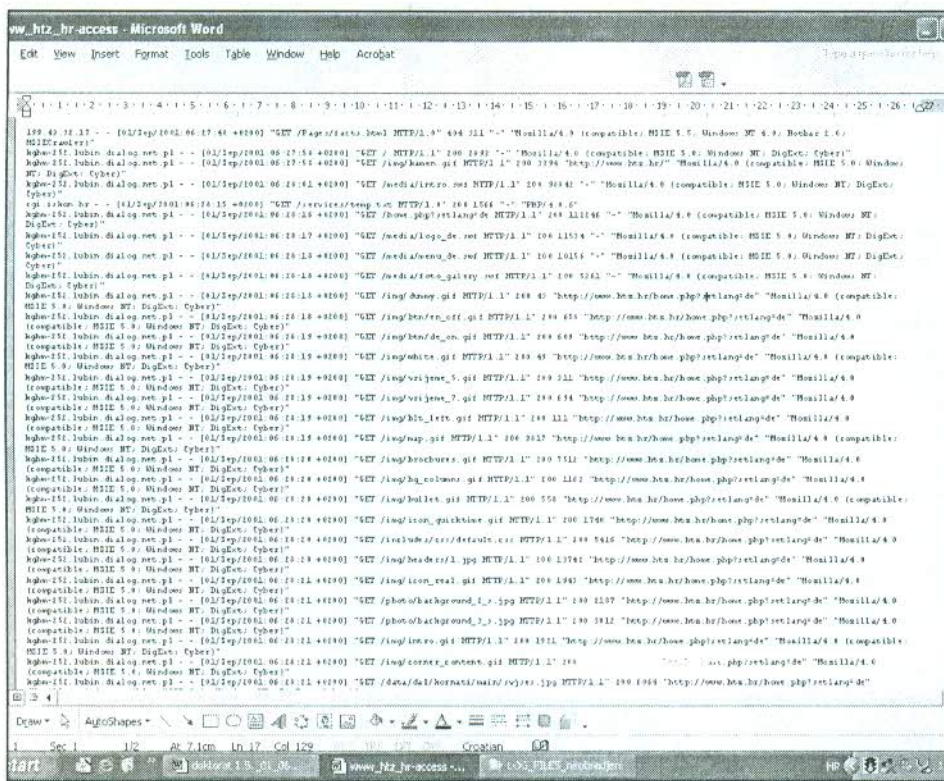
Što se tiče zastupljenosti elemenata multimedije, početnu se stranicu ne smije bespotrebno memorijski "nafilati" nepotrebnim Kb (kilobajtima), budući još uvijek o njima ovisi brzina učitavanja istih (npr. ponuditi html i flash rješenje). U budućnosti će se i to ograničenje sasvim sigurno adekvatno riješiti, no do tada to ograničenje ipak valja imati na umu. Upravo se iz toga razloga preporuča korištenje teksta i slike kao za ovaj entitet najadekvatnijih elemenata multimedije. Glede varijabli kojima bi sadržaj kao "podentitet" Index Entiteta trebao omogućavati pristup do svih modelom raspoloživih informacija, navodi se samo jedno ograničenje odnosno zahtjev, a to je njegova informacijsko-sadržajna ukupnost koja će biti definirana kroz ostale entitete. Izradi ostalih entiteta pristupa se na isti način kao i izradi prikazanog Index entiteta.

2. LOG DATOTEKA KAO TEMELJ ANALIZE TOKOVA KORIŠTENJA WEB STRANICA KAO MMIM-A I PREDUVJET ZA POBOLJŠANJE POSLOVNOG REZULTATA

2.1. LOG DATOTEKA - njene karakteristike i značaj

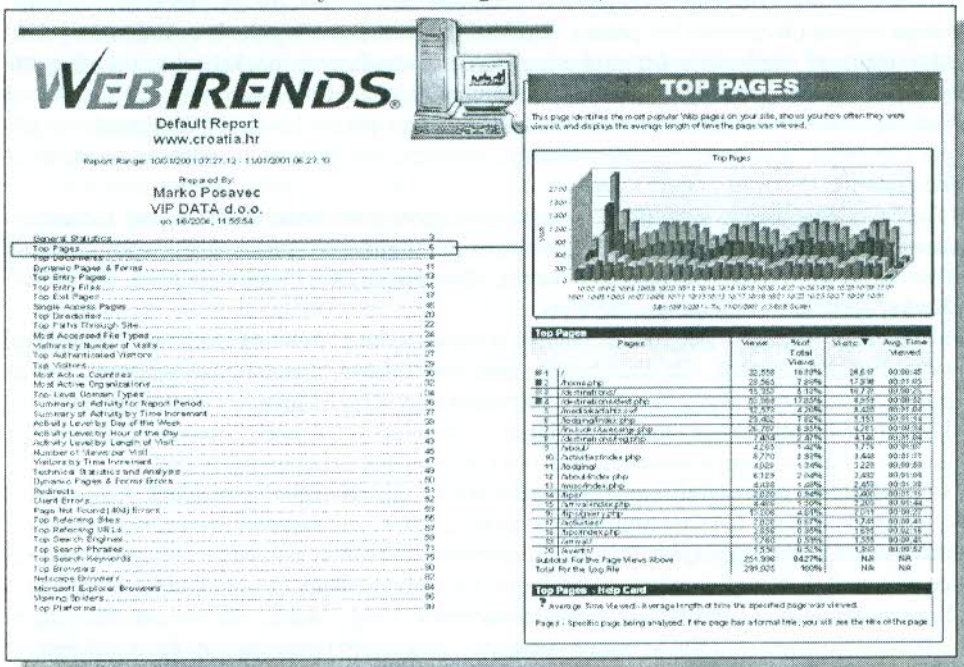
Osnovne karakteristike Log datoteke očituju se stohastički i neobrađenom zapisu kojeg kreira web server i koji sadrži podatke o tome tko je posjetio Web čvor, kada, te koje je stranice (a u sklopu njih i slike koje su njihov sastavni dio) posjetio/učitao (vidi sliku 8.1). Međutim, zapis log datoteke može se na puno višem nivou analizirati i iskoristiti, što će se kroz daljnji tijek ove radnje pokušati objasniti i obrazložiti.

Slika 2.1.: Stohastički i neobrađeni zapis log datoteke



Uglavnom svaki ozbiljan web host (davatelj web usluga u širem smislu) daje serverske logove na uvid webmasterima određenih web čvorova kako bi mogli analizirati svoj promet (ili daje već obrađene log datoteke, tj. gotovu statistiku). Log datoteke koristi velik broj programa za analizu prometa, no postoje i aspekti promatranja cijele problematike analize prema kojima bi se značaj iste trebao postaviti na jedan viši nivo. Što to znači? Obrazloženje bi se trebalo očitovati kroz cijelu ovu radnju, a najkraće rečeno trebalo bi log datoteku što više iskoristiti u smislu nastojanja produbljivanja njene analize u statističko-semantičkom smislu i osigurati programski vođenu regulaciju web čvora što nije nimalo lagani zadatak. Naime, uglavnom sve analize loga datoteka svode se na korištenje jednog od postojećih, analitičkih programa (Webtrends, Webalizer) što se vidi na slijedećoj slici, čije je rezultat svakodnevnan detaljan grafičko-tablični ispis svega onoga što se na serveru događa.

Slika 2.2.:Primjer obrađene log datoteke (Webtrends statistika)



Izvor: "Web trends default report" uz doradu autora

No najčešće je tu i kraj "priče". Zašto je to tako? Odgovor će se pokušati dati kroz sljedeće činjenice:

- Web analize jako su opširne
- Web analize nisu usko profilirane i prilagođene struci (dakle pojedinim vrstama web čvorova)
- Poslovni je svijet iz dana u dan sve dinamičniji, menadžeri nemaju vremena svaki dan izgubiti nekoliko sati samo za proučavanje statističkih izvještaja, da bi onda donosili odluke što i kako dalje sa web čvorom.
- «Web analize», dobro su početno zamišljene, ali su sa vremenom i razvojem tehnologije postale nedostatne.

2.2. Log datoteka kao izvor podataka i temelj za analizu

Digitalno okruženje omogućava primjenu novih metoda za praćenje i analizu korištenja digitalnih izvora. Uz pomoć programske podrške moguće je automatski pratiti sve aktivnosti ili transakcije ostvarene s računalnim sustavom poslužitelja, tzv. **log file analysis**. Analiza log datoteka prvotno je bila namijenjena administratorima poslužitelja za praćenje opterećenosti sustava, no s vremenom su istu počeli koristiti i analitičari raznih profila i potreba. Osnovni elementi analize pristupa log datoteke sadrže odgovore na sljedeća pitanja:

- **Tko?** – identifikacija korisnika; u većini slučajeva identificira se IP adresa računala s kojeg korisnik pristupa. Kada je pristupanje regulirano preko korisničkog računa (username i password), moguća je osobna identifikacija korisnika i njegovih aktivnosti.
- **Kada?** – datum i vrijeme pristupanja
- **Što?** – identifikacija korištenih materijala-web stranica, i načina korištenja materijala (statistička analiza). Za standardno praćenje postoji programska podrška, no za potrebe ovoga rada ista je unapređena i "podignuta" na viši nivo.
- **Kako?** - identifikacija načina korištenja informacija, tj web stranica (semantička analiza). Za praćenje je potrebna dodatna i unapređena programska podrška.
 - Prilikom prikupljanja podataka ne smiju se previdjeti postojeći nedostaci i nesavršenosti te se na njih mora ukazati što će se i u nastavku ovoga rada učiniti:
- Aktivnosti korisnika registriraju se do trenutka dobivanja podataka iz baze, stoga ne znamo na koji su način dalje korišteni, npr. čitaju li korisnici direktno na zaslonu?
- Pretraživanja se u predmemoriji (cache memory) mogu zadržavati po nekoliko dana. Ukoliko korisnik u tom vremenu ponovno zatraži isti sadržaj on će ga dobiti iz predmemorije pa taj pristup na poslužitelju neće biti registriran.
- Dolaze li registrirani zahtjevi od jedne ili više osoba. Nije uvijek moguće razlučiti dolaze li višekratni upiti za isti dokument od jednog korisnika ili od više njih. Individualnog korisnika ne može se identificirati preko IP adrese ukoliko nema stalnu vezu na Internet, a i kada ju ima vezu sistematično se ne može pratiti pristupanje iste osobe ukoliko ona koristi više računala. Isto tako više osoba može koristiti isto računalo, dakle za različite upite bilježi se ista IP adresa pa također ne dobivamo točne

podatke o korištenju. Treba napomenuti da je praćenje određenog korisnika moguće u slučajevima kada je korištenje određenog web čvora uvjetovano unosom "logirajem", dakle upisom korisničkog imena (engl.username) i lozinke (engl. password).

2.3. Stohastika podataka u sklopu log datoteke

Log datoteka kao tekstualna (txt) ASCII datoteka predstavlja računalnu odnosno podatkovnu sliku tokova korištenja web čvora od strane određenog broja korisnika. Ta slika u svakom je momentu drugačija što znači da istu možemo okarakterizirati kao dinamičnu. Ovisno o veličini i «popularnosti» Web čvora, ta slika može biti brže ili sporije dinamična, što naravno ovisi o broju korisnika koji «surfa» po Web čvoru, veličini Web čvora, aktivnostima korisnika na web čvoru i vremenima korisnika provedenim tražeći informaciju na određenom Web čvoru. Svaka aktivnost svih različitih korisnika bilježi se u Log datoteci. Buduće stanje zapisa u log datoteci ne može se unaprijed predvidjeti upravo zbog činjenice da se nikakva buduća interakcija korisnika sa Web čvorom ne može predvidjeti. Stoga se može tvrditi da je log datoteka skoro apsolutno stohastička. Isto tako, zapisi u log datoteci nisu poredani po nikakvoj unaprijed definiranoj strukturi, a po formatu se razlikuju od servera do servera. Zapis u log datoteci vrlo je promijenljiv, a uvijek se prati i pohranjuje prema unaprijed definiranim vremenskim intervalima (sat, dan, tjedan, mjesec,...) što je najčešće determinirano veličinom web čvora i intenzitetu «surfanja» po njemu o čemu posljedično ovisi i sama veličina log datoteke i dinamika promijena u njoj. Za obradu gore navedenih zapisa-podataka (slika neobrađene log datoteke) potrebno je imati program koji će prepoznati format log datoteke, pročitati sve potrebne podatke, strukturirati iste i pohraniti ih u odgovarajuću bazu (moguće je i direktno upisivanje «log-a» u bazu podataka). Svi tako sakupljeni podaci služe za daljne analize, prema kojima je i nastala ideja izrade cjelokupnog DDWS modela.

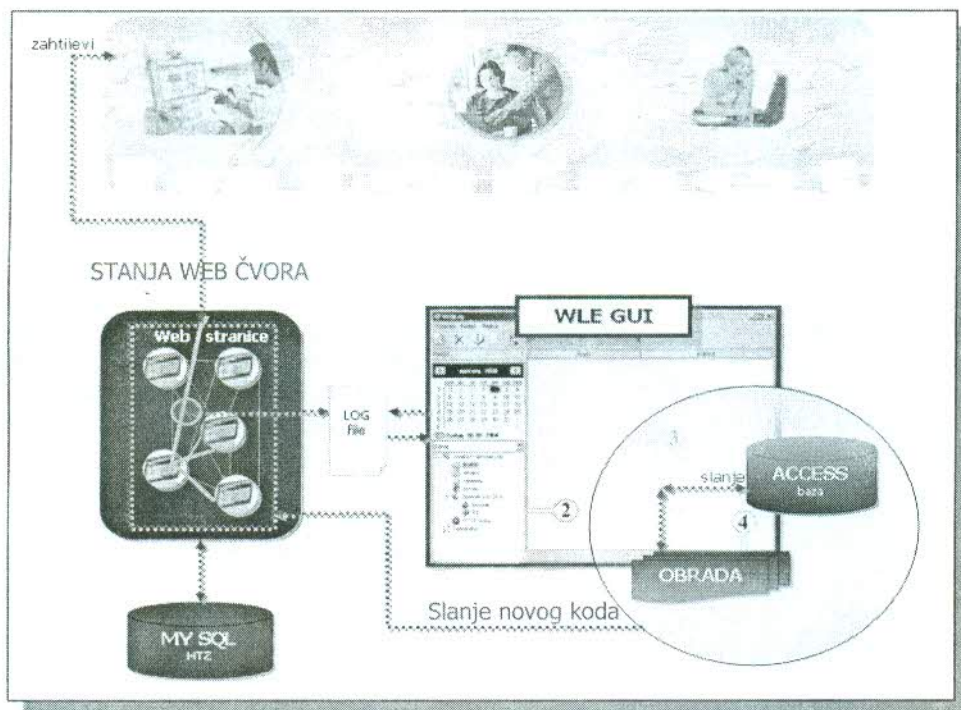
2.4. Potreba strukturiranja i stvaranja modela podataka log datoteke

Kao što je već i okvirno navedeno da bi se nad log datotekom svrhovita kvantitativno-kvalitativna analiza nužno je postojeće "sirove" podatke transformirati i prilagoditi za analizu. Transformacija i prilagodba podataka za analizu, pripremanje i pisanje analitičkog algoritma i nije tako trivijalan posao. Stoga je prije samoga dubljeg ulaska u problematiku i njeno rješavanje, potrebno izvršeno istraživanje radi upoznavanja sa svjetskim dostignućima u navedenom problemskom području. U nastavku ovoga poglavlja prikazuje se ideja DDWS modela, te WLE programskog rješenja kao njegovog implementacijskog dijela.

3. TEORIJSKE I PRAKTIČNE POSTAVKE KVANTITATIVNO KVALITATIVNE ANALIZE MMIM-A

Kada govorimo o kvantitativno kvalitativnoj analizi, onda u suštini mislimo na statističko semantičku analizu log datoteke kao izvora podataka. U skladu sa takvim pristupom i shvaćanjem u nastavku se objašnjavaju pristupi i principi statističke i semantičke analize⁵ log datoteke u sklopu DDWS modela. Osnovni prikaz DDWS modela prikazan je na slici 3.1.

Slika 3.1.: DDWS model



⁵ Jane Greenberg, Stuart Sutton, D Grant Campbell: « Metadata: A fundamental component of the Semantic Web», Bulletin of the American Society for Information Science and Technology. Silver Spring: Apr/May 2003. Vol. 29, Iss. 4;

3.1. DDWS Model – Statistička analiza

Jedan od najvažnijih razloga zbog kojeg se može oblikovati DDWS model je analiza. Najčešće se koriste statističke i semantičke analize.

Statistička analiza za DDWS se radi iz obrađenih i spremljenih podataka u bazi podataka, a u sklopu web čvor-a sastoji se od slijedećih podanaliza:

- analiza posjećenosti stranica
- analiza prometa slika
- analiza prometa servera
- analiza vremena (učestalosti) u kojem je prikazan promet po vremenu (od 0 – 24 sata)
- analiza prometa browsera i operativnih sustava (OS-a)

3.1.1. DDWS Model – Analiza posjećenosti stranica

Posjećenost stranica dobiva se tako da se filtriraju svi podaci iz baze koja u sebi sadrži adrese stranica odabranog web čvor-a koje je pojedini korisnik posjetio. Adrese su ograničene samo na one koje u sebi sadrže određene ekstenzije stranica (.php, .asp, .htm i dr.). Svaka adresa vezana je uz datum i vrijeme «klik» pojedinog korisnika na istu, tako da se može napraviti izvršiti potpuna analiza posjećenosti stranica u određenom vremenu.

3.1.2. DDWS Model – Analiza prometa po slikama

Promet po slikama dobiva se na sličan način kao i posjećenost po stranici s razlikom da se u ovom slučaju uzimaju sve adrese slika koje se učitavaju pri posjetu pojedinog dijela web čvor-a. Uz svaku adresu slike također je vezan datum i vrijeme, tako da se može izvršiti analiza prometa po slikama u određenom vremenu. Ponovo se napominje da se slike najčešće učitavaju zbog svoga smještaja u sklopu pojedinih web stranica čiji su sastavni dio.

3.1.3. DDWS Model – Analiza prometa po serverima

Promet po serverima dobiva se na sličan način kao i posjećenost po stranici s razlikom da se u ovom slučaju uzimaju adrese servera ili TCP/IP adresa korisnika koji je posjetio web čvor. Također uz svaku adresu servera ili TCP/IP adresu vezan je datum i vrijeme kada se to dogodilo. Na kraju moguće je dobiti analizu prometa po serverima unutar određenog vremenskog intervala.

3.1.4. DDWS Model – Analiza prometa u vremenskom intervalu

Promet u vremenskom intervalu (po satima) dobiva se tako da se uzmu sve adrese stranica odabranog site-a koje sadrže ekstenzije (.php, .asp, .htm i dr.) i razvrstaju po

vremenskom intervalu, te se na kraju može izvući analiza prometa u vremenskim intervalima (u ovom slučaju po satima) unutar određenog perioda (datuma).

3.1.5. DDWS Model – Analiza prometa po browser-u i OS-u

Promet po browser-u i OS-u dobiva se tako da se filtriraju svi podaci koji u sebi sadrže naziv i verziju browser-a sa kojim je korisnik pretraživao na odabranom web čvor-u i naziv i verzija korištenog operativnog sustava kojim pristupa web čvor-u. I ti su podaci također vezani uz datum i vrijeme, sa čime se može dobiti analiza u određenom vremenskom periodu. Isto tako, mogu se napraviti i razne druge analize koje ovise o podacima log datoteke odnosno podešenosti web servera da prikaže željeni format log datoteke. Log datoteka se «puni» podacima koji se dobivaju interakcijom korisnika sa web čvor-om što je već prije spomenuto. Neke od takvih analiza su: promet po količini učitanih i poslanih bajtova svakog podatka (reda) iz log datoteke, promet po zastupljenosti država iz kojih se najčešće ide na odabrani Web čvor i slično.

3.2. DDWS Model – Semantička analiza

Jedna od dodatnih analiza za DDWS i ne manje važnih je semantička analiza web čvor-a. Ona u sebi sadrži sve podatke potrebne za kreiranje navigacijske karte u određenom periodu, odnosno navigacijskih puteva kojima je korisnik «surfao» da bi došao do željene informacije na određenom web čvor-u. Pri tome je korištena teorija grafova. Za semantičku analizu primjenjen je usmjereni graf sa oznakama puteva. Prema njima se mogu saznati neke od bitnih informacija:

- od strane korisnika najčešće korišteni «pathovi», grupirani da bise dobila njihova količina,
 - najčešća dubina gore spomenutih «pathova»,
 - poletne točke u «pathu» od kojih najčešće korisnici kreću, kao i u kojoj točki najčešće završe,
 - zastupljenost određene adrese web čvor-a u bilo kojem dijelu «patha»
- Pri tome moraju se uvesti pojedina pravila prema kojima je moguće dobiti željenu analizu:
- da bi dobili «path» moramo imati najmanje dvije točke
 - ako nema niti jednog «patha» za taj server odnosno ako je to prvi «path» servera onda se otvara novi «path»
 - ako zadnji «path» tog servera sadrži adresu stranice onda se otvara novi «path», to znači da se dogodio ciklus (iz teorije grafova početna i krajnja točka su iste). Inače dodaje se adresa stranice zadnjem «path-u» tog servera.
 - novi put (path) se otvara ukoliko se promijeni naziv host-a slijedeći gornja pravila
 - ako HTTP status nije 400 ili veći (što znači da je došlo do greške pri učitavanju stranice)

Ostvarenje same DDWS ideje moguće je ostvariti izradom programskog rješenja koje bi svojom multifunkcionalnošću osiguralo ne samo statističko semantičku analizu web čvora kao MMIM-a, već i njegovu permanentnu dinamiku, odnosno prilagodbu trenutnim zahtjevima tržištu potražnje. Sa tim je i ciljem izrađeno WLE programsko rješenje koje se u nastavku opisuje.

4. TIJEK IZRADE WLE PROGRAMA ZA ANALIZU LOG DATOTEKE

Prije sam izrade programa odabrana je baza podataka MS Access (budući je bila dostatna za eksperimentalni tijek istraživanja, no za primjenu u praksi za složenije modele preporuča se korištenje MySQL-a, MSSQL-a, Oracle, PostgreSQL,...) i programski jezik C# iz Visual studia .NET 2003⁶ čime je moguće ispuniti naše zahtjeve i kriterije.

Projektiran je DBMS, kreirane su tablice, atributi i međusobne relacijske veze. Nakon što je odabran programski alat i baza podataka u koju ćemo pohranjivati podatke iz log datoteke, napravljene su osnovne klase za povezivanje programa i baze podataka. Klasa za bazu podataka sadrži većinom funkcije za spremanje i čitanje podataka. Izrađena je klasa za što točniju obradu loga, pri tome poštujući kriterije koji su ranije postavljeni.

Za kreiranje formata log datoteke prema želji uvedena je specifikacija tagova kao pomoć WLE programu sa slijedećim opisom:

- %S naziv servera
- %d dan
- %Mmjesec
- %Y godina
- %h sat
- %m minuta
- %s sekunda
- %E metoda (GET,PUT,POST)
- %e URL put
- %q URL query
- %c status servera
- %f naziv hosta
- %B browser i OS
- %j bilo koji drugi znak (razmak i sl.)

Prije samog slaganja tagova potrebno je znati format log datoteke. Osmišljeno je i grafičko sučelje i ostalih dijelova programa da bi korisnik mogao što lakše primijeniti program. Nakon toga bilo je potrebno kreirati funkcije za dobivanje potrebnih statističkih analiza u određenom vremenskom periodu.

⁶ M.Stiefel, R.J.Oberg:»Application development using C# and .NET«, Prentice Hall PTR, NJ 07458, 2002.

Za dobivanje semantičke analize izrađena je klasa kojom je moguće razvrstati obradene podatke po serverima i definirati navigacijske putove kojima je pojedini korisnik posjetio određene web stranice u sklopu web čvora.

Nakon toga kreiran je prozor i logika za dinamičku promjenu stranica. Pri tom su iskorišteni pohranjeni podaci kako bi se definirala pravila prema kojima će se dinamika odvijati, kao i «zamjenski kod» koji se treba umetnuti da bi se mijenjao sadržaj stranica. Sa time je ispunjen zamišljeni koncept DDWS modela.

5. ANALIZA STRUKTURIRANIH PODATAKA I UTJECAJ NA WEB ČVOR

Kada govorimo o analizi strukturiranih podataka i povratnom utjecaju na Web čvor, tada već polako zaziremo u samu funkcionalnost DDWS modela. DDWS model mora osigurati obavljanje dva zadatka:

- obraditi trenutno stanje posjećenosti web čvor-a
- izmjene web čvor-a (povratni utjecaj)

Obrada trenutnog stanja se radi prema skupljenim podacima unutar nekog perioda u kojem su se događale interakcije korisnika sa web čvor-om. Ti podaci se nalaze u log datotekama web servera čiji se zapis može događati u određenim vremenskim intervalima te se po formatu razlikuje od servera do servera. Uobičajeni intervali su: vremenski (sat), dnevno, tjedno, mjesečno ili prema dostignutoj veličini datoteke.

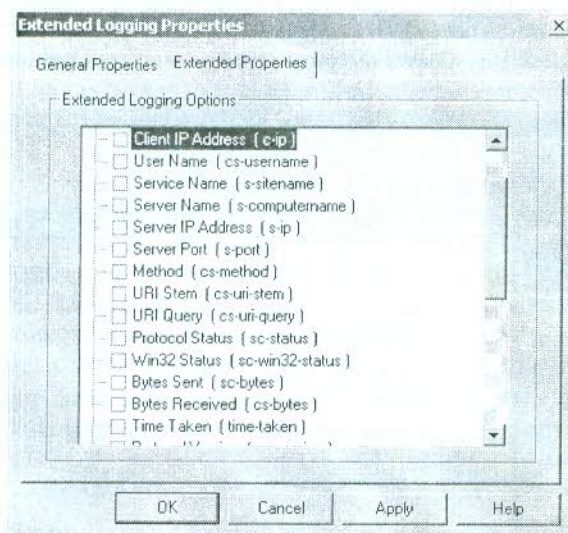
Osnovni format log datoteke prema W3C standardu je sljedeći:

- datum i vrijeme
- TCP/IP adresa korisnika
- naziv servera
- metoda
- web adresa (link)
- web query (link query)
- status servera
- host korisnika
- (browser i operativni sustav)

Neki serveri imaju mogućnost zapisa i još nekih dodatnih parametara kao na primjer: početak i kraj session-a korisnika, količina primljenih i poslanih bajtova stranice itd. Poredak parametara formata ovisi o web serveru (slika 5.1: Parametri na IIS serveru).

Za obradu navedenih podataka potrebno je imati program koji će prepoznati format log datoteke, pročitati sve potrebne podatke i pohraniti ih u odgovarajuću bazu. Svi tako sakupljeni podaci služe za daljnje analize, prema kojima se izvršava drugi zadatak DDWS-a.

Slika 5.1: Parametri na IIS serveru



Izmjena web čvor-a obuhvaća promjene sadržaja točno odabranih stranica web čvor-a, a u skladu sa trenutnim stanjem korisničkih interakcija. Odluke o izmjenama se donose na osnovu određenih kriterija: posjećenosti stranica, učestalosti servera (korisnika), učestalosti navigacijskih putova i slično. Ti podaci se dobiju iz gore navedenog programa i baze podataka. Promjenu sadržaja može napraviti sam osoba koja je zadužena za takove poslove bez pomoći specijaliziranog programa za dinamičke izmjene ili se taj postupak može djelomično automatizirati primjenom takvog programa. Potpuna automatizacija je ograničena prepoznavanjem raznih web jezika i logike kojom su «pisane» web stranice. Budući da se dinamika želi realizirati na već postojećim web čvor-ovima, problem bi bio ispravno tumačiti kod web programera i odrediti mjesta izmjene. Zato se umjesto toga radi automatizacija u kojoj se osobi zaduženoj za izvršavanje izmjena (update) ostavlja da sam odluči, ili se to učini prema njegovim zahtjevima - koji dio koda i na koji način treba promijeniti, ali se same izmjene koda i stranice rade preko specijaliziranog programa (WLE program).

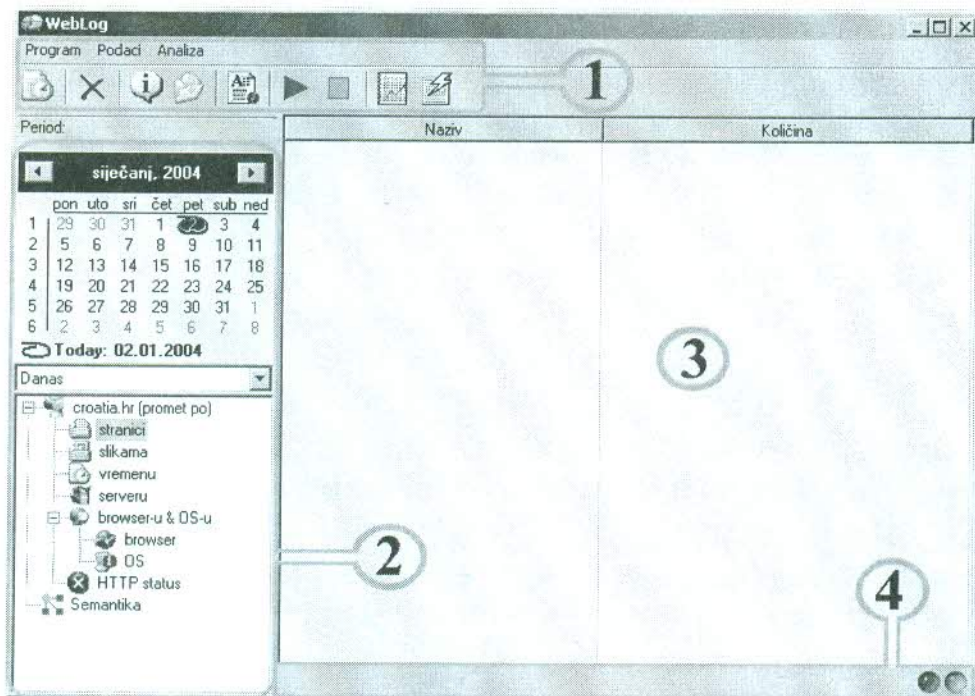
6. WLE – KRATKI OPIS DIJELOVA PROGRAMA

Program je podijeljen u nekoliko cjelina:

- glavni dio programa
- dio za pregled i ispis izvještaja
- dio za podešavanje parametara

- dio za podešavanje parametara
- dio za dinamičku promjenu

Slika 6.1: WLE - Glavni dio programa (početno grafičko sučelje)



Zbog ograničenosti rada i same važnosti u nastavku se ukratko opisuje samo glavni dio WLE programskog rješenja, budući bi opisivanje cijelog programskog rješenja i njegovih modula i podmodula zahtijevalo znatno sustavniji i opširniji pristup čime bi se okviri ovoga rada znatno premašili.

6.1. WLE Program – glavni dio

Glavni dio programa sastoji se od 4 dijela:

- izbornika i trake s ikonama (toolbar-a)
- lijevog panela (koji u sebi sadrži kalendar, filter i kontrolu sa područjima koja se mogu analizirati)
- desnog panela (list kontrola koja služi za prezentaciju podataka)
- statusne trake (koja pokazuje da li je pokrenuta obrada podataka)

Kako bi se objasnio DDWS model provesti će se analize – istraživanja, koja proizlaze iz mogućnosti i karakteristika WLE programa.

WLE program kao implementacija i najvažniji dio DDWS modela ima višedimenzionalnu ulogu u smislu prepoznavanja formata log datoteke, njene obrade na osnovi koje postoji mogućnost raznih analiza u određenim vremenskim intervalima, te na koncu zadavanja naloga za dinamičku izmjenu web stranice u sklopu web čvora.

Prvi zadatak WLE programa jest prepoznavanje formata loga datoteke i njena obrada. Kao što je već i u više navrata u ovome radu napisano, log datoteka je vrlo neuredna i stohastična sadržaja. Isto tako vrlo je mala da se njen sadržaj dvaput ili više puta ponovi, što znači da svako pokretanje WLE programa daje različite rezultate, odnosno uvid u trenutno stanje i aktivnosti na strani tržišta potražnje čiju sliku log datoteka predstavlja. Dakle, u kontekstu takova razmišljanja izvršiti istraživačku analizu znači provesti cjelokupni WLE proces nad određenom log datotekom. Za potrebe ovoga rada prikupljeni su podaci-log datoteke Web čvora HTZ-a, od mjeseca srpnja, 2003. godine (dakle 31 log datoteka). Prosječan dnevni «log» iznosi cca. 600.000 slogova što predstavlja veliku količinu podataka i odražava reprezentativnu sliku aktivnosti tj. interesa tržišta potražnje.

LITERATURA

1. E.G.Abels, M.D.White, K.Hahn:"*A user based design process for Web Sites*", Internet Research: Electronic Networking Applications and Policy, vol.8.,n.1., 1998., str.39-48.
2. J.M.Artz:"*A top down methodology for building corporate Web applications*", Internet Research:Electronic Networking Applications and Policy, vol.6.,n.2/3., 1996., str.64-74.
3. I.Avelini-Holjevac:"*Kontroling – upravljanje poslovnim rezultatom*", Hotelijerski fakultet Opatija, 1998.
4. I.Avelini-Holjevac:"*Total Quality Management u hotelijerstvu i turizmu*", Hotelska kuća '96., str 5., Hotelijerski fakultet Opatija, 1996
5. F.Beard, R.L.Olsen:"*Webmasters as mass media gatekeepers: a qualitative exploratory study*", Internet Research:Electronic Networking Applications and Policy, vol.9.,n.3., 1999. str. 201-211.
6. H.Davulcu, J.Freire, M.Kifer, I.V.Ramakrishnan.: "*A layered Architecture for Quering Dynamic Web Content*", Bell Laboratories.
7. J. Degner:"*What is good hypertext writing?*", 1994.
8. J.A. DeVito:"*Human Communication: The basic course*", Harper Collins College Publishers, New York, 1993.
9. J.Eaton, J.Smithers, S.Curran:"*This is it: A manager's Guide to Information Technology*", Philip Allan Publishers Limited, New Jersey, 1982.
10. J.H. Ellsworth & M.V. Ellsworth:"*Marketing on the Internet*", John Wiley & Sons, Inc.,Toronto, 1997.
11. D. M Ewalt: "*The next Web*", InformationWeek. Manhasset: Oct 14, 2002. , Iss. 910
12. G.Gillespie:"*Too much info hazardous to your health*", The Institute, vol 21., n.6., june, 1997.
13. W. Gitt:"*Information-die dritte Grundgröße neben Materie und Energie*", Siemens Zeitschrift, No.4, 4-9, 1989.
14. J. Gordon: "*Multimedia motion*", Folio. Stamford: Aug 2003. Vol. 32, Iss. 8
15. P.Gralla:"*How Internet Works*", Ziff-David Press, Emeryville, CA, 1997.
16. M.J.Granger, D.L. Schroeder:"*Integrating the Internet into the business environment*", Internet Research: Electronic Networking Applications and Policy, vol.6.,n., 2/3., 1996. str.,85-89.
17. D.I. Hawkins, R.J. Best., Coney, K.A.: "*Consumer behavior*", Irvin, 1995.

18. Ž.Hutinski, M.Šimunić: "Status zaštite podataka u hrvatskom turizmu", Ekonomija i ekologija u funkciji razvoja turizma, Hotelijerski fakultet Opatija, 1999.
19. Ž.Hutinski, M.Šimunić, M.Randić: "Information Model(ing) of Web pages for hotel presentation on Internet", DAAAM Symposium, Vienna, 1999. str. 205-206.
20. Izvještaj Comiteta AMA za definicije, Journal of Marketing, vol XIII. No 2.
21. M.Kifer, G.Lausen, J.Wu: " Logical Foundations of Object-Oriented and frame based Languages", Department of Computer Science, Suny, Stony Brook, NY,1994. , (Accepted to Journal of ACM)
22. T.Kesić: "Marketinška komunikacija", Mate, Zagreb, 1997.
23. S.-E. Kim, Thomas Shaw, Helmut Schneider: "Web site design benchmarking within industry groups" Internet Research, Bradford, 2003. Vol. 13, Iss. 1
24. A.C. Luther: "Authoring Interactive Multimedia", Academic Press, Cambridge, 1994.
25. W.May: " Information Extraction from the web with Florid", Dagstuhl Seminar, Declarative Dana on the Web, Septembar, 1999.
26. M. Randić, M. Šimunić: "Prezentacija hotela na Internetu", Hotelijerski fakultet Opatija, Hotelska kuća '98., str. 261-272, 1998.
27. M. Randić, M. Šimunić, P.Knežević: "Modelling structure and dynamic behaviour of Web pages augmented with dynamic HTML", DAAAM Symposium, Vienna, 1998. str. 413-414.
28. M. Šimunić: "Suvremena komunikacija s tržištem – modeli prezentacije hotela na Internetu", Fakultet za turistički i hotelski menadžment, Opatija, 2000. – magistarski rad
29. M.Stiefel, R.J.Oberg: «Application development using C# and .NET», Prentice Hall PTR, NJ 07458, 2002.
30. F. Williams: "The New Communications", 3rd ed. Belmont, CA. Wadsworth, 1992

WWW izvori i reference

- America Institut of Stress - <http://www.stress.org>
- Internet Society - <http://www.isoc.org/internet/history/>
- Sveučilište u Freiburgu (F-logic, FLORID) - <http://www.informatik.uni-freiburg.de/~dbis/florid/>
- World Wide Web Consortium (W3C) - <http://www.w3.org/>
- Webmaster Magazine - <http://www.web-master.com>
- The Internet Engineering Task Force - <http://www.ietf.org> – (<ftp://ftp.rfc-editor.org/in-notes/rfc3508.txt>, <ftp://ftp.rfc-editor.org/in-notes/rfc1738.txt>, <ftp://ftp.rfc-editor.org/in-notes/rfc2616.txt> , <ftp://ftp.rfc-editor.org/in-notes/rfc1630.txt>, <ftp://ftp.rfc-editor.org/in-notes/std/std5.txt>, <ftp://ftp.rfc-editor.org/in-notes/rfc1866.txt>,