

aLive!
- sistem za inteligentno prosleđivanje pitanja -
aLive!
- Intelligent Question Forwarding System -

Stanko Nikolić, Bojan Furlan, Pavle Josipović

Elektrotehnički fakultet Beograd

Sadržaj – U ovom dokumentu je opisan sistem koji korisnicima interneta omogućava brz, efikasan i kvalitetan način za dobijanje odgovora na pitanja.

Abstract – This document contains description of the system which gives fast, effective and high-quality way for getting answers.

1. UVOD

Živimo u svetu velikog tehnološkog razvoja. Proteklih godina postignut je ogroman napredak na polju dostupnosti informacija, pre svega u svetu internet pretraživača. Prvi put u svojoj istoriji, ljudi su u mogućnosti da u izuzetno kratkom vremenu dođu do novih saznanja. Ipak, sva rešenja koja trenutno postoje, poseduju određene slabosti.

Uprkos visokom stepenu razvoja veštačke inteligencije, ljudski mozak je i dalje superiornije i jeftinije rešenje kada je u pitanju rezonovanje ljudskog ponašanja i konverzacije. Ni jedno veštačko rešenje ne može da zameni osećaj razumevanja koji možemo videti u očima prijatelja sa kojim razgovaramo.

Sa druge strane, daleko je od humanog eksploatacija čoveka za visoko zahtevne intelektualne poslove. Svaki čovek bi trebalo da radi ono za šta je najbolji i ono što najviše voli. Ukoliko želite logičan odgovor na neko konkretno pitanje, najbolje je da, umesto lutanja po moru informacija koje nam internet nesumnjivo pruža, pitate osobu koja će vas razumeti i ponuditi Vam kratak i jasan odgovor.

Osnovna ideja projekta aLive! je da se pomogne ljudima da dođu do prave osobe, bez gubitka dragocenog vremena.

Računari mogu da, u dovoljnoj meri, odrede oblast na koju se određeno pitanje odnosi. Takođe, na osnovu aktivnosti korisnika mogu da zaključe koliko je nekom korisniku bliska određena oblast. Ono što računari nisu u mogućnosti da odrade je shvatanje suštine pitanja i pružanje adekvatnog, ažurnog i konciznog odgovora. Taj deo, koji je od najvećeg značaja za korisnika, delegira se pravoj osobi.

Ovim je stvoren sistem u kome čovek i mašina saraduju na pravi i najefikasniji način. Posao koji je zadržan na strani ljudi je distribuiran na veliki broj korisnika, i kao takav ne predstavlja opterećenje pojedinačnom korisniku. Upravljanje celim sistemom, tj. pronalaženje prave osobe za određeni zadatak predstavlja zahtevan i dosadan posao koji je prepušten mašini. Ovakvom podelom, odnosno, prepuštanjem relativno malog ali izuzetno značajnog dela posla čoveku, postiže se veoma visoka efikasnost.

2. PRIJAVLJIVANJE NA SISTEM

Za prijavljivanje na sistem, neophodno je da korisnik poseduje Windows Live ID. Ovim načinom autentifikacije obezbeđeno je da postojeći korisnici Windows Live servisa mogu krajnje jednostavno da postanu korisnici ovog sistema. Takođe, sva briga oko administracije korisničkih naloga delegirana je trećoj strani.

Prilikom prvog prijavljivanja na sistem, od korisnika se zahteva da popune formular. Time se dobija početna informacija o njihovim interesovanjima, bitna za rad glavnog algoritma.

3. POSTAVLJANJE PITANJA

Ključna funkcionalnost sistema predstavlja mogućnost korisnika da pitanje uputi sistemu, koji će obezbediti kvalitetan odgovor. Radi lakše upotrebe sistema od strane korisnika, radna površina je jednostavna, tako da korisnik lako uočava prostor predviđen za unos pitanja.

4. PROSLEĐIVANJE PITANJA

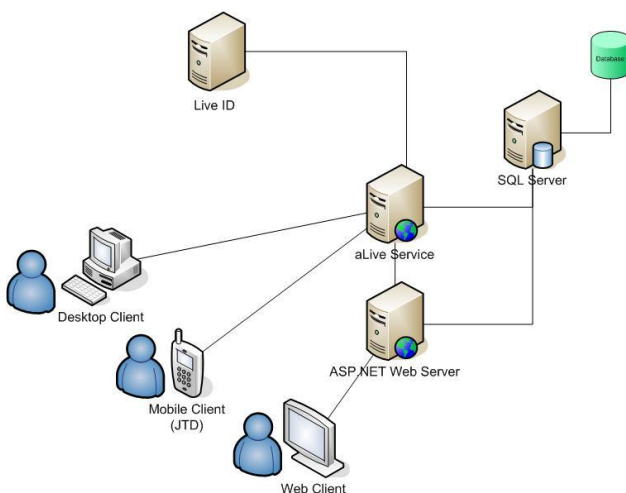
Pošto je neki korisnik postavio pitanje sistemu, sistem bira određeni broj korisnika koji su najkompetentniji i prosleđuje im pitanje kao instant poruku. Sistem izbora je detaljno objašnjen u sekciji *algoritam odlučivanja*. Korisnici koji su dobili prosleđeno pitanje, u mogućnosti su da odbiju, ako trenutno nisu u mogućnosti da odgovore, ili da odgovore na pitanje, popunjavanjem textbox-a i klikom na dugme za prosledjivanje odgovora. Korisnik je u mogućnosti da bira da li će taj odgovor poslati anonimno ili želi da pošalje svoj potpis zajedno sa njim.

5. ODGOVARANJE NA PITANJE

Kada korisnik odgovori na pitanje, sistem prosleđuje odgovor onom korisniku koji ga je postavio. Korisnik prima odgovor u vidu instant poruke. On može da oceni kvalitet odgovora i može označiti da ne želi da više dobija odgovore na to pitanje.

6. ARHITEKTURA SISTEMA

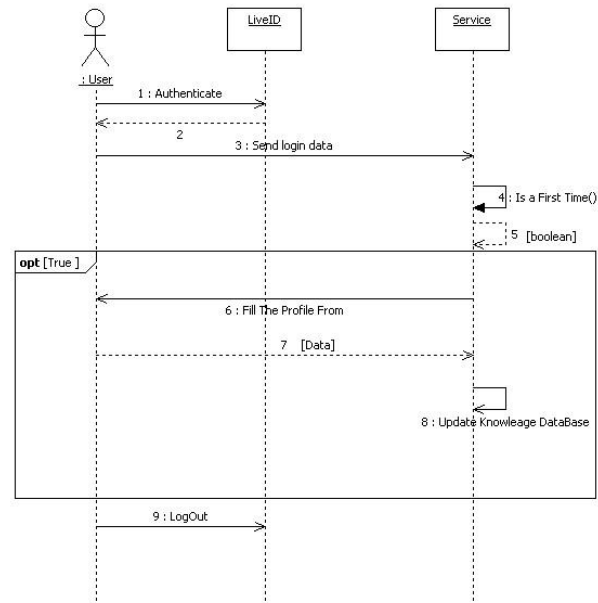
Centralnu komponentu sistema predstavlja aLive! Web servis kome mogu pristupiti različite klijentske aplikacije. Autentikacija se obavlja pomoću LiveID Web servisa, dok se registracija i pristup različitim uslugama obavlja pomoću ASP.NET web aplikacije.



Slika 1 - arhitektura sistema

7. SCENARIO KORIŠĆENJA

Na slici 2 prikazan je scenario korišćenja sistema.

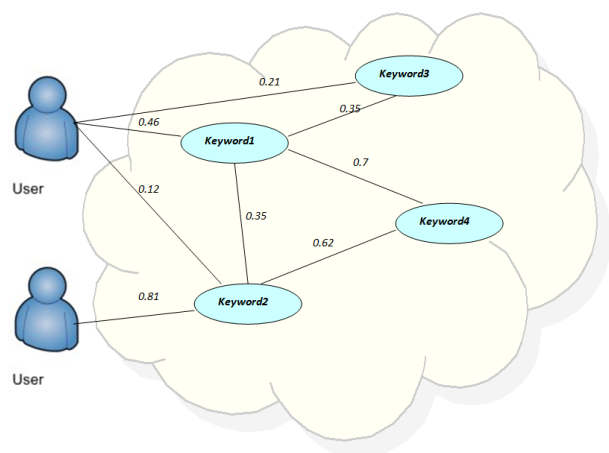


Slika 2 - scenario korišćenja

8. STRUKTURA BAZE ZNANJA

Predstava znanja zasniva se na vezama između ključnih reči (koje će biti prepoznate u pitanju i odgovoru) i korisnika i veza između samih ključnih reči. Sve veze u sistem sadrže težine čije su vrednosti skalirane između 0 i 1, tj. predstavljaju verovatnoće. Da bi sistem pravilno funkcionisao neophodno je da vrednosti veza budu pažljivo podešene.

U sistemu postoje dve vrste veza : „ključna reč – ključna reč“ i „korisnik – ključna reč“.

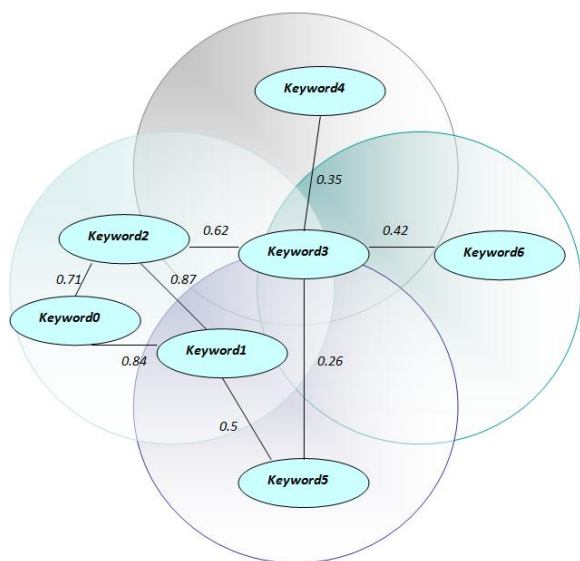


Slika 3 - struktura baze znanja

Podešavanje veza se vrši kroz dva slučaja korišćenja (PostingQuestion i ReceiveAnswer). Na ovaj način sistem uči o interesovanjima određenih korisnika, kao i o međusobnim vezama između ključnih reči.

Veza „korisnik – ključna reč“ svojom težinom predstavlja korisnikova interesovanja prema određenim ključnim rečima, pa samim tim i prema određenim oblastima znanja što će biti objašnjeno u daljem tekstu.

Veza „ključna reč – ključna reč“ je neophodna da bi sistem imao mogućnost da odabere korisnike koji u trenutku postavljanja pitanja nisu direktno vezani za datu ključnu reč koja se nalazi u postavljenom pitanju, ali imaju veliku verovatnoću vezivanja sa ostalim ključnim rečima bliskim ovoj ključnoj reči (Pod izrazom „blisko“ samtramo veliku verovatnoću vezivanja). Na ovaj način u stabilnom stanju sistema dinamički će se formirati oblasti oko određenog skupa reči, koje će konvergirati sa određenom težinom ka tom skupu. Takođe, određene ključne reči mogu konvergirati ka više skupova obzirom da se neke oblasti mogu preklapati.



Slika 4 - formiranje abstraktnih oblasti

Konačni cilj je da postavljeno pitanje dođe do korisnika koji imaju najveću verovatnoću vezivanja sa ključnim rečima iz pitanja ili drugim ključnim rečima koje su vezane sa datim ključnim rečima, tj. pripadaju istoj oblasti znanja.

Da bi baza znanja bila u konzistentnom stanju ažuriraju se dva globalna brojača, koji predstavljaju :

- Globalni maksimum neskaliiranih vrednosti veza korisnik - ključna reč
- Globalni maksimum neskaliiranih vrednosti veza ključna reč - ključna reč.

Skalirane vrednosti se dobijaju kada se neskaliirana vrednost neke veze podeli sa brojačem koje je zadužen za tu vrstu veze.

Oba ova brojača imaju svoju maksimalnu vrednost. Kada brojač dostigne maksimum, svim vezama (onim za koje je određeni brojač zadužen), neskaliirane vrednosti se dele sa 2 zajedno sa brojačem, na taj način, njihove diskretne vrednosti bivaju prepolovljene, ali međusobni odnosi ostaju jednaki. Posle ovakve intervencije, novoažurirane veze imaju veću težinu a ostale zastarevaju, tj. gube na značaju ukoliko se ne koriste istim intenzitetom.

9. ALGORITAM ODLUČIVANJA

Srž ovog sistema je inteligencija koja određuje kako sistem da odabere najkompetentije korisnike za odgovor na postavljeno pitanje od strane drugog korisnika sistema.

Ona se zasniva na posebnom algoritmu veštačke inteligencije osmišljenom za ovaj konkretan problem, zasnovanom na algoritmima mašinskog učenja.

Uvedene su sledeći pojmovi:

skup1 – skup svih obrađenih ključnih reči

skup2 – skup svih novih ključnih reči koje su u vezi sa rečima iz Skupa 1

V_k – sračunata vrednost ključne reči

V_k' – vrednost ključne reči pomoću koje je algoritam našao datu ključnu reč

L_k – težina veze pomoću koje je algoritam našao datu ključnu reč

S – stepen specifičnosti pitanja za navedene ključne reči u pitanju

r – razdaljina, tj. korak iteracije u algoritmu.

Pronalaženje ključnih reči za rangiranje korisnika:

Za pronalaženje ključnih reči za rangiranje korisnika koristi se sledeći algoritam:

1. Korisnik pre postavljanja pitanja određuje stepen specifičnosti njegovog pitanja za navedene ključne reči, tj. koeficijent koliko će algoritam široko ići u dohvatanju novih ključnih reči (u daljem tekstu koeficijent **S**).
2. Sve ključne reči vezane za dato pitanje inicijalno formiraju **Skup1**.
3. Dohvataju se sve nove ključne reči vezane sa ključnim rečima iz Skupa1 čije su težine veze veće od koeficijenta **S**. Od novoizabranih reči se formira **Skup2**. U slučaju da je Skup2 prazan algoritam se prekida.
4. Za sve ključne reči iz Skupa2 računa se vrednost po formuli:

$$V_k = \frac{L_k * V'_k}{2^r}$$

Napomena: Inicijalno $V'_k = 1, r = 0$. U slučaju da se neka ključna reč može naći pomoću više reči iz prethodnog skupa, uzima se najveća vrednost. Uvođenjem koeficijenta $1/2^r$ postiže se da težina opada sa povećanjem rastojanja, tako da reči udaljenije od ključnih reči korišćenih u pitanju će učestvovati sa manjim koeficijentom.

5. Sve ključne reči čija je sračunata vrednost veća od koeficijenta **S** premeštaju se iz Skup2 u Skup1.
6. Proces se ponavlja od tačke 3.

Kreiranje rang liste korisnika:

Algoritam ispituje težinu veza između svih ulogovanih korisnika i ključnih reči iz Skupa1, izuzev korisnika koji je postavio pitanje.

Za svakog ulogovanog korisnika računa se suma :

$$V_u = \sum V_k * L_u$$

V_k – sračunata vrednost ključne reči za dato pitanje

L_u – težina veze korisnik - ključna reč

V_u – sračunata vrednost za datog korisnika

Na osnovu V_u sistem određuje **N** najboljih korisnika kojima prosleđuje pitanje.

Na kraju terba povećati vezu između korisnika koj je postavio pitanje i ključnih reči i vezu između ključnih reči u pitanju, za diskretnu vrednost.

10. PRIJEM ODGOVORA

Korisnik koji je postavio pitanje (i dobio odgovor), potrebno je da oceni kvalitet odgovora, ali pre toga potrebno je odrediti njegovu kompetentnost.

Da bi se odredila kompetentnost korisnika koj daje ocenu moramo znati njegov koeficijent povezanosti sa ključnim rečima u pitanju (u daljem tekstu K_p). K_p se dobija na taj način što se izračuna aritmetička sredina svih njegovih veza sa ključnim rečima vezanim za postavljeno pitanje.

Dalje, korisniku su ponuđene četiri ocene:

- OK

Potrebno je povećati težinu veza između korisnika koj je odgovorio i svih ključnih reči u pitanju, za diskretnu vrednost proporcionalnu K_p .

- Neutralno

Ne vrši se modifikacija težina veza.

- Loše

Potrebno je smanjiti težinu veza između korisnika koj je odgovorio i svih ključnih reči u pitanju, za diskretnu vrednost proporcionalnu K_p .

- Zloupotreba sistema

Povećava se *abuse* parametar za datog korisnika.

Na kraju, algoritam povećava vrednost veze između ključnih reči u odgovoru i korisnika koji je odgovorio, kao i vezu između samih ključnih reči, za diskretnu vrednost.

11. ZAKLJUČAK

Opisani sistem pruža korisnicima izuzetno visok kvalitet usluge. Sistem ne postavlja ograničenja u pogledu govornog područja, oblasti primene, geografske udaljenosti ili kompjuterskih resurasa na strani klijenta. Korisnici su oslobođeni zamornog rada koji se odnosi na pronalaženje prave osobe za dobijanje odgovora, i taj posao je prepušten serveru. Implementacijom sistema u obliku instant messenger servisa, dobija se na brzini. Međutim, kada je u pitanju shvatanje pitanja, upravo prepuštanje tog posla živoj osobi podiže kvalitet ovog sistema u odnosu na postojeća rešenja.

Karakteristika	Internet pretraživač	Internet Forum	Instant Messenger	aLive!
Prirodna formulacija pitanja	-	+	+	+
Shvatanje konteksta pitanja	+	-	-	+
Pronalaženje korisnika odgovarajućeg profila	-	-	-	+
Prosleđivanje pitanja korisnicima	-	-	+	+
Shvatanje suštine pitanja	-	+	+	+
Odgovaranje	-	+	+	+
Brzo dobijanje odgovora	+	-	+	+

Slika 5 - poređenje sa postojećim sistemima

LITERATURA

- [1] MSDN Library - <http://msdn2.microsoft.com/en-us/library/default.aspx>
- [2] Microsoft .NET Framework 3.0 Resources - <http://www.netfx3.com>
- [3] Microsoft Gadgets - <http://microsoftgadgets.com/>
- [4] Windows Live SDK - <http://msdn2.microsoft.com/en-us/library/bb264574.aspx>
- [5]. **Russell, Stuart J.** / **Norvig, Peter.** Artificial Intelligence - A Modern Approach. 1995.
- [6]. **Weiss, Gerhard.** Multiagent Systems - A Modern Approach to Distributed Modern Approach to Artificial Intelligence. 1999.
- [7]. **Bigus, Joseph** / **Bigus, Jennifer.** Constructing Intelligente Agents Using Java. 2001.