

Rangiranje Univerziteta u Beogradu zasnovano na Webometrics metodologiji i preporuke za postizanje boljeg ranga

Miloš R. Todorović, Jelica Ž. Protić

Sadržaj — Webometrics, koji meri prisutnost univerziteta na Internetu, prema opisanoj metodi, pokazao je drastičan pad prisutnosti Univerziteta u Beogradu u poslednjem merenju. U cilju postizanja boljih rezultata i pronalaženja odgovarajućih preporuka akademskim članicama Univerziteta, sproveden je niz eksperimenata i merenja koja su omogućila uvid u bolju sliku trenutnog stanja i moguće promene koje bi dovele do bolje pozicije rangiranja.

Gljučne reči — Internet, merenje prisutnosti, rangiranje univerziteta, web, webometrics.

I. UVOD

VRHUNSKI univerziteti objavljuju milione strana koje proizvode desetine odseka i službi, stotine istraživačkih timova i hiljade naučnika. Snažno prisustvo na Internetu informiše nas o širokom spektru faktora koji su u jasnoj korelaciji sa opštim kvalitetom institucije: široka dostupnost kompjuterskih izvora, globalna Internet pismenost, politike koje promovisu demokratiju i slobodu izražavanja, konkurencija za međunarodnu vidljivost i podrška inicijativama za otvoren pristup.

Cilj ovog rada je upoznavanje sa pojmom rangiranja prisutnosti univerziteta na Internetu, metodologijom i problemima u analizi. Data je analiza rezultata rangiranja Univerziteta u Beogradu, kratak pregled analize domena akademskih članica Univerziteta i faktora koji utiču na ukupni rang, uočenih problema i mogućnosti za njihovo rešavanje.

Na kraju rada dat je zaključak o neophodnim akcijama i mogućim promenama koje bi dovele do bolje pozicije rangiranja Univerziteta.

II. WEBOMETRICS

Rangiranje univerziteta na Internetu na svetskom nivou ("World Universities ranking on the Web") je inicijativa Cybermetrics Lab, istraživačke grupe Centra za humanističke i društvene nauke (Centro de Ciencias Humanas y Sociales, CCHS), koja je deo Nacionalnog saveta za istraživanje (National Research Council, CSIC), najvećeg javnog istraživačkog tela u Španiji (www.csic.es).

M. R. Todorović, Univerzitet u Beogradu, Elektrotehnički fakultet, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija; (telefon: 381-11-3218-320; faks: 381-11-3248681; e-mail: mt@etf.bg.ac.rs).

J. Ž. Protić, Univerzitet u Beogradu, Elektrotehnički fakultet, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (telefon: 381-11-3218-456; faks: 381-11-3248681; e-mail: jeca@sezampro.rs)

Cybermetrics Lab se bavi kvantitativnom analizom Interneta i sadržaja na Internetu, posebno onih koji su vezani za procese generisanja naučnog znanja i naučne komunikacije.

Webometrics [1]-[2] rangiranje meri količinu, vidljivost i uticaj veb stranica koje objavljuju univerziteti. Cilj rangiranja je da se promoviše objavljivanje na Internetu, i oceni posvećenost elektronskoj distribuciji, kao i borba protiv akademske digitalne podele koja je evidentna čak i među univerzitetima u razvijenim zemljama.

Direktna ciljna grupa kojoj su namenjeni ovi rezultati rangiranja su uprave univerziteta, kreatori veb stranica fakulteta i nastavnici i istraživači kao glavni akteri pripreme i postavljanja adekvatnih sadržaja.

III. MERENJE PRISUTNOSTI

Webometrics je studija koja se zasniva na kvantitativnom merenju i koja indirektno obuhvata i kvalitativni aspekt strukture, korišćenja izvora informacija i tehnologija na Internetu (WWW), oslanjajući se na bibliometrijski i informetrijski pristup [3]. Webometrics uključuje analizu veza (linkova), analizu citiranja na Internetu, ocenjivanje od strane pretraživača i čisto opisne studije [4]-[5].

Veliki problem u analizi veza, koja je osnova merenja prisutnosti sadržaja na Internetu, je to što se Internet konstantno menja i čini se da je u stalnoj ekspanziji, tako da zaključci vebometrijske analize mogu vrlo brzo postati zastareli. Niz višegodišnjih i intenzivnih ispitivanja nad sajtovima univerziteta u Australiji, Novom Zelandu i Velikoj Britaniji ukazali su na ovaj problem. Merenja su pokazala da su se ispitivani univerzitetski veb sajtovi stabilizovali po veličini od 2001 godine, posle nekoliko godina brzog rasta [6]. Poređenjem linkova između veb sajtova iz godine u godinu pokazalo se da je ova stabilizacija po veličini prikrila promene u pojedinačnim vezama, ali je zaključeno da bi tipične kvantitativne analize i pored toga mogle da imaju rok trajanja i više godina [7].

Značajan udeo u vebometrijskom istraživanju dat je kroz ocenu komercijalnih pretraživača [8]. Istraživanja u okviru razvoja algoritama za prikupljanje informacija, kao i načina kako se koriste pretraživači za traženje informacija nisu deo vebometrije. Obzirom da komercijalni pretraživači samo delimično pokrivaju sadržaj na Internetu [9], postavlja se pitanje da li postoji

neki bitan faktor koji utiče na merenje prisutnosti? Ovo pitanje je ključno, jer su pretraživači posrednici između korisnika i veb sadržaja što im daje značajnu ekonomsku moć u novoj sferi ekonomije na Internetu [10]-[11].

IV. METODOLOGIJA RANGIRANJA

Osnovna hipoteza je da broj linkova [12] koji ciljaju akademski domen jeste proporcionalan obimu istraživanja tj. produktivnosti zajednice institucija, na nivou univerziteta [13], katedri [14], istraživačkih grupa [15], i/ili samih naučnika pojedinačno [16].

Jedinica za analizu je domen institucije, a u ovom konkretnom slučaju domen univerziteta.

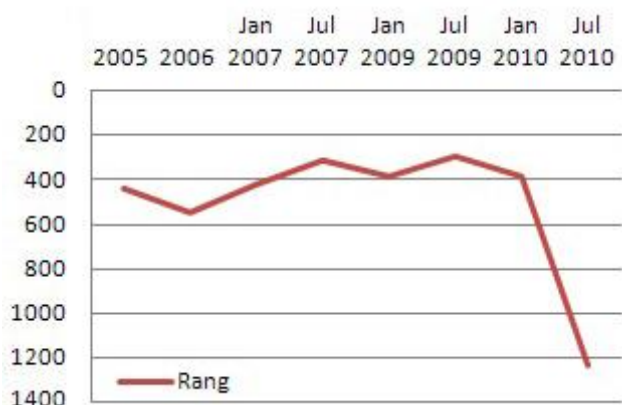
Rangiranje se vrši kombinovanjem grupe indikatora koji mere različite aspekte [17]. Almind & Ingwersen predložili su prvi veb indikator, Faktor uticaja na Internetu (Web Impact Factor-WIF) [18], koji se zasniva na analizi linkova kombinujući broj spoljnih dolaznih linkova (external inlinks) sa brojem stranica na veb sajtu i to u razmeri 1:1 između vidljivosti i veličine [19]. Ova razmera se koristi za rangiranje uz dodatak dva nova indikatora o komponenti veličine: broj dokumenata (rich files) i broj publikacija koje se nalaze u Google Scholar [20] bazi podataka.

Formula rangiranja na osnovu ova četiri faktora: Veličina (S) (broj indeksiranih strana od strane pretraživača), Vidljivost (V) (broj jedinstvenih spoljnih linkova), Rich Files (R) (broj fajlova sa bogatim sadržajem: pdf, ps, doc, ppt) i Scholar (Sc) (broj radova i citata dobijenih preko Google Scholar servisa); glasi:

$$\text{Rang} = 0.5 * V + 0.25 * S + 0.15 * R + 0.15 * Sc$$

V. RANGIRANJE UNIVERZITETA U BEOGRADU

Univerzitet u Beogradu se uvek dobro rangirao među svetskim univerzitetima. Od kad se sprovede rangiranja, pa do poslednjeg rangiranja nalazio se u prvih 500, a u pojedinim periodima i u prvih 300 univerziteta u svetu.



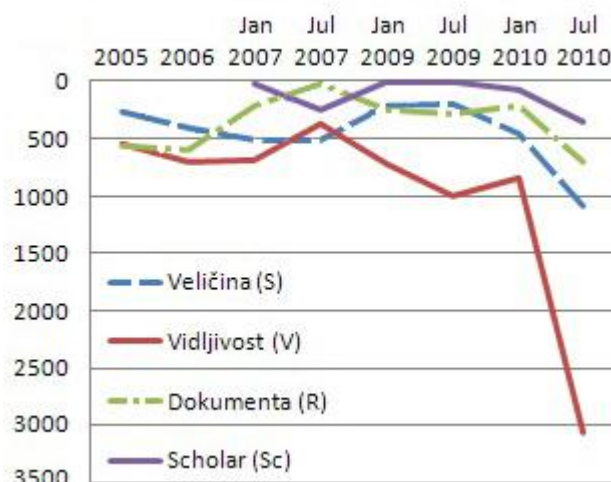
Sl. 1. Prikaz ogromnog pada pozicije rangiranja Univerziteta u Beogradu pri poslednjem merenju

Rangiranje se radi od 2004. godine i prve godine su za indikatore uzeti samo veličina i vidljivost domena i tu Univerzitet u Beogradu svakako nije mogao da bude konkurentan univerzitetima sa daleko boljom infrastrukturom i praksom postavljanja sadržaja na

Internet.

Sledeće dve godine malo je izmenjen algoritam i dodat je faktor broja dokumenta (Rich Files), čime se Univerzitet u Beogradu izborio za ulazak u prvih 500. Ovaj faktor je dosta unazadio neke univerzitete, a nekima dosta pomogao. U savremenim zemljama postojala je već praksa da se dokumenti postavljaju na biblioteke koje se nalaze na Internetu, a čijim se sadržajima obično pristupalo preko zaštićenog dela. Sa druge strane, analizom je utvrđeno da na našim prostorima ljudi više vole da postavljaju dokumenta, nego da kreiraju veb stranice sa istim sadržajem.

Slična stvar je usledila i kasnije sa uvođenjem Google Scholar faktora, 2007 godine. Ne postojanje mesta gde bi se postavili sadržaji sa konferencija, radovi, projekti i sl., učinili su ove materijale javno dostupnim i time lako indeksiranim od strane pretraživača. To objašnjava rang među prvih 10 kod ovog faktora, prve godine i kasnije 2009. godine, što je značajno popravilo ukupni rang.



Sl. 2. Prikaz rangiranja Univerziteta u Beogradu po godinama prema faktorima.

Na slici 2, prikazan je veliki pad vidljivosti koji je posledica prestanka važenja .yu domena i pojave višestrukih domena u okviru fakultetskih sajtova članica Univerziteta od uvođenja novog .rs domena, čime je izgubljen značajan broj linkova koji su vodili ka sajtovima Univerziteta u Beogradu. Ove činjenice su direktno uticale i na ostala tri faktora, ali u manjoj meri, obzirom na količinu materijala, dokumenata i radova objavljenih poslednjih godina.

VI. ANALIZA DOMENA ČLANICA UNIVERZITETA

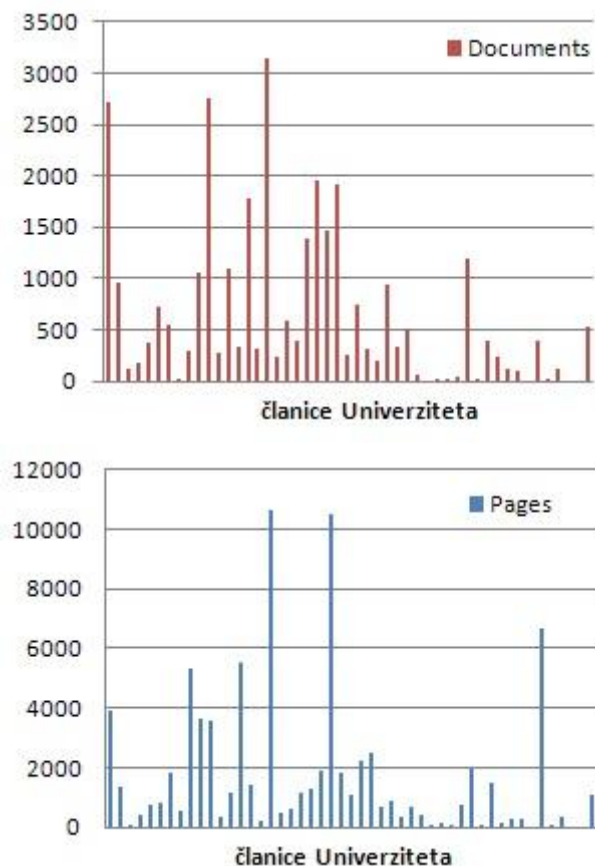
Detaljnou analizom domena i veb sajtova članica Univerziteta u Beogradu, došlo se do zanimljivih rezultata.

Analizom vidljivosti dokumenata na domenima članica, a sprovodeći identičan postupak prikupljanja podataka kao webometrics ranking, utvrđeno je da postoje velika odstupanja u količini pronađenih dokumenata, kao i poražavajuća činjenica da mnoge institucije nemaju praksu postavljanja dokumenata na Internet, što se posebno odnosi na institute i razvojne centre.

Mnogi urednici sajtova nisu svesni činjenice da sadržaj

koji su postavili nije dostupan pretraživačima za indeksiranje iz jednostavnih tehničkih razloga: ćirilčni nazivi fajlova, predugački nazivi i loši linkovi, loša organizacija menija, nepovezanost među sajtovima i sl.

Loša tehnička rešenja sajta, nedostatak upotrebe meta tagova, loša organizovanost informacija, nedostatak povezanosti ka i među podsajtovima univerzitetskih članica su samo neki od uzročnika slabog indeksiranja stranica od strane pretraživača.



Sl. 3. Grafici prikazuju velike razlike u broju pronađenih dokumenata i indeksiranih stranica na domenima članica Univerziteta u Beogradu, koje su posledica loših tehničkih rešenja, loše organizacije i drugih faktora.

Na velikom broju sajtova uočeni su linkovi sa .yu domena, koji ne postoji već duže vreme, što govori o velikoj neažurnosti pojedinih sajtova. Ovde se naročito radi o engleskim verzijama sajtova koje su skoro potpuno zanemarene. Čak na dva od tri svetski priznata sajta (<http://univ.cc>, <http://www.bulter.nl/universities/>) na kojima se nalazi popis svih univerziteta u svetu se Univerzitet u Beogradu još uvek nalazi sa .yu linkom.

Neažurnost i nepostojanje stranica na engleskom sajtu indirektno utiču na smanjenje broja dolaznih linkova. Na pojedinim sajtovima zabeleženo je ispod 3% ili čak ni jedna stranica na engleskom.

Prilikom pregleda sajtova, uočeno je da su pojedini sajtovi redizajnirani ili su u fazi radova. Razvojne verzije se nalaze na alternativnim adresama, postoje redirekcije linkova, dosta stranica nema sadržaj ili su isključene funkcionalnosti i što je najbitnije ne postoje veze sa starim

sadržajem, čime se značajno gube informacije.

Dodatni problem, koji se pojavio uvođenjem .rs domena je pojava alternativnih domena. Pojedine članice su zakupile dodatni domen, kao alternativni naziv postojećem domenu poddomena .bg.ac.rs. i na taj način se distancirale od Univerziteta u Beogradu. Čak i u slučaju kad oba domena ukazuju na isti sadržaj ili postoji redirekcija, uvođenjem alternativnih domena gube se spoljni dolazni linkovi, smanjuje broj indeksiranih stranica i dokumenata i stvara nepovezanost institucija članica Univerziteta, što je upravo jedan od elemenata na koje webometrics merenja i analize rangiranja žele da ukažu.

VII. PREPORUKE ZA BOLJE RANGIRANJE

Sledeće preporuke imaju za svrhu da pomognu univerzitetima i drugim institucijama (instituti i razvojni centri) da postignu adekvatno prisustvo na Internetu.

Veliku prisutnost na Internetu je moguće ostvariti samo uz napore i zalaganja veće grupe autora. Najbolji način da se ovaj savet primeni je ohrabriranjem i podržavanjem velikog broja naučnika, istraživača i studenata da budu potencijalni autori ili urednici sajtova predmeta, katedri i projekata, a posebno stimulisati veb sajtove konferencija, naučnih društava i njihovih publikacija, elektronske biltene i sl.

Veoma je važno da se izbegne menjanje institucionalnog domena, jer bi to stvorilo zabunu i imalo loš efekat na stopu vidljivosti. Alternativne ili domene ogledala treba izbegavati čak i kada preusmeravaju na pravi domen.

Resursi koji nisu dostupni u elektronskoj formi mogu se lako konvertovati u veb stranice. Mnogi univerziteti imaju dugo beležene podatke o aktivnostima koji se mogu objavljivati na veb sajtovima.

Za očekivati je da sajtovi akademskih institucija poseduju linkove od strane „prirodnih“ partnera: institucija na lokalnom nivou ili iz regiona, veb direktorijuma iz sličnih organizacija, portala koji pokrivaju direktne teme institucije, ličnih stranica kolega ili partnera.

Verzije na drugim jezicima, posebno na engleskom, su obavezne ne samo za glavne stranice već i za odabrane delove/sekcije i posebno za naučna dokumenta. Ovo je naročito bitno, jer je engleski jezik upravo onaj jezik koji pretraživači razmeju.

Iako je HTML standardni format za veb stranice, nekad je bolje da se upotrebe rich file formati kao što su Adobe Acrobat pdf ili MS Word doc jer oni omogućuju bolju distribuciju dokumenata. Prilikom kreiranja ovih dokumenata potrebno je obezbediti da nazivi fajlova budu čitljivi na Internetu, kao i popuniti polja za naziv autora, vreme kreiranja, autorska prava i sl.

Prilikom izrade treba izbegavati glomazne menije za navigaciju koji se zasnivaju na Flash, Javi ili JavaScript-u. Previše duboka podela i kompleksno razgranavanje sadržaja mogu takođe blokirati pristup robotima pretraživača; izbaciti iz upotrebe uvodne stranice sajtova sa odabirom jezika.

Loše organizovane baze podataka, pa čak i veoma dinamične strane mogu biti nevidljive za neke pretraživače, pa je u ovim slučajevima bolje koristiti

direktorijume ili statične strane umesto postojećeg rešenja ili kao još jednu dodatnu opciju.

Postoji čitava nauka koja se bavi optimizacijom sajtova tako da budu što prilagođeniji pretraživačima (SEO-Search Engine Optimisation), jer od toga, koliko su pretraživači u stanju da indeksiraju stranice sajta, najviše zavisi vidljivost sadržaja, odnosno prisutnost sajta.

Upotreba relevantnih naslova i deskriptivnih meta tagova može da poveća vidljivost stranica. Praćenjem vrlo jednostavnih saveta i pravila koja su data na sajtovima pretraživača i servisa može se značajno doprineti vidljivosti i poboljšati kvalitet sajta.

Održavanje kopija starog ili arhivskog materijala na sajtu bi trebalo da bude obavezno. Ponekad se važne informacije izgube kada se sajt redizajnira ili jednostavno ažurira, i ne postoji način da se lako vrate stranice koje se u tom procesu izgube.

VIII. ZAKLJUČAK

Webometric indikatori služe da se pokaže posvećenost institucije objavljivanju na Internetu. Ukoliko je pozicija jedne institucije na Internetu ispod pozicije koju su očekivali u skladu sa svojim akademskim postignućima, trebalo bi da razmotre veb politiku, kako bi promovisali znatno povećanje u količini i kvalitetu svojih elektronskih publikacija.

Analizom rezultata rangiranja i pojedinačnom analizom domena članica Univerziteta u Beogradu, došlo se do zaključka da je neophodno promeniti politiku i odnos prema objavljivanju sadržaja na Internetu, kao i uvesti niz pravilnika o korišćenju domena, načinu objavljivanja sadržaja, odgovornosti institucija za sadržaj itd.

Ukoliko se žele postići bolji rezultati na svetskoj sceni i rang listi, neophodno je povećati obim i kvalitet sadržaja na stranom jeziku ili jezicima, kao i objavljivanje materijala koji bi bio interesantan i posetiocima van domaćeg govornog područja.

Neophodno je izvršiti dodatne analize kvaliteta sadržaja sajtova članica, utvrditi i postaviti određene minimume i standarde koje bi sajtovi morali da ispunjavaju.

Na kraju, neophodna je saradnja svih strana kako bi se podigla svest o značaju informacija objavljenih na Internetu i popravio kvalitet sadržaja, što bi vremenom trebalo da dovede i do pojave posebnog tela čija bi uloga bila praćenje kvaliteta, ukazivanje na nepravilnosti, pomoć članicama, stimulacija i ohrabrivanje grupa koje bi održavale i kreirale nove sadržaje na Internetu.

LITERATURA

- [1] Sangam, S.L. and Munavalli, S.B. IT literature: "A webometrics study". *Library progress (international)*, vol. 24 (1), 2004. p. 15-21.2.
- [2] T.C. Almind and P. Ingwersen, Informetric analyses on the World Wide Web: "Methodological approaches to Webometrics", *Journal of Documentation* 53(4) (1997) 404-426.
- [3] Goswami, Prashant. "Webometric Study of Universities Websites of Uttar Pradesh". MPHIL Dissertation, Vikram University, Ujjain. 2006-07.
- [4] L. Vaughan and D. Shaw, Bibliographic and Web citations: "What is the difference?" *Journal of the American Society for Information Science and Technology* 54(14) (2003) 1313-1322.

- [5] L. Vaughan and D. Shaw, "Web citation data for impact assessment: A comparison of four science disciplines", *Journal of the American Society for Information Science & Technology* 56(10) (2005).
- [6] N. Payne and M. Thelwall, "A longitudinal study of academic webs: Growth and stabilization", *Scientometrics* 71(3) (2007)
- [7] N. Payne "A longitudinal study of academic Web links: Identifying and explaining change". (University Of Wolverhampton, Wolverhampton, 2007)
- [8] J.Bar-Ilan, "The use of Web search engines in information science research". *Annual Review of Information Science and Technology* 38 (2004) 231-288.
- [9] S. Lawrence and C.L. Giles, "Accessibility of information on the web", *Nature* 400 (1999) 107-109.
- [10] L. Introna and H. Nissenbaum, "Shaping the web: Why the politics of search engines matters", *The Information Society* 16(3) (2000).
- [11] E. Van Couvering, "New media? The political economy of Internet search engines". In *Annual Conference of the International Association of Media & Communications Researchers*. (Porto Alegre, Brazil, 2004)
- [12] K. Kousha and M. Thelwall "Motivations for URL citations to open access library and information science articles", *Scientometrics* 68(3) (2006) 501-517.
- [13] M. Thelwall "Extracting macroscopic information from web links", *Journal of American Society for Information Science and Technology* 52(13) (2001) 1157-1168.
- [14] O. Thomas and P. Willet "Webometric analysis of departments of librarianship and information science", *Journal of Information Science* 26(6) (2000) 421-428.
- [15] F. Barjak and M. Thelwall "A statistical analysis of the web presences of European life sciences research teams", *Journal of the American Society for Information Science and Technology* (2008)
- [16] F. Barjak, X. Li, and M. Thelwall "Which factors explain the web impact of scientists' personal home pages?" *Journal of the American Society for Information Science and Technology* (2007).
- [17] L. Björneborn and P. Ingwersen "Toward a basic framework for webometrics", *Journal of the American Society for Information Science and Technology* 55(14) (2004) 1216-1227.
- [18] P. Ingwersen "The calculation of Web Impact Factors", *Journal of Documentation* 54(2) (1998) 236-243.
- [19] B. Cronin "Bibliometrics and beyond: some thoughts on web-based citation analysis", *Journal of Information Science* 27(1) (2001)
- [20] K. Kousha and M. Thelwall "Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis" *Journal of the American Society for Information Science and Technology* 58(7) (2007) 1055 -1065.

ABSTRACT

Academic institution's footprint is based on evaluation of the scholarly activities of academic organizations, scholars and researchers, measured by web indicators in three large groups: activity (web publication), impact (visibility) and usage (visits and visitors). Webometrics, which measures Web presence of Universities, according to the described methodology, has shown noticeable decrease of presence of the University of Belgrade during last year. In order to achieve higher ranking of our University through recommendations to the member faculties regarding their Web sites, we have performed quantitative experiments and measuring that provided a clear picture of the present state and possible changes that would lead to ranking improvement.

RANKING OF THE UNIVERSITY OF BELGRADE BASED ON WEBOMETRICS AND RECOMMENDATIONS FOR ACHIEVING BETTER RANK

Miloš Todorović
Jelica Protić