

Тема: **Имплементација система за решавање словних CAPTCHA у Java технологији**

Наставник: **Александар Цветковић**

Циљ мастер рада је израда пакета који решава словне CAPTCHA у Java технологији. Употреба CAPTCHA приликом израде web site-а има прилично једноставно објашњење, наиме, треба склонити садржину web site-а од евентуалних посета web bot-ова, spider-а, crawler-а, harvester-а и сличног softwer-а који се користи приликом индексирања web страница. Идеја је користећи методе компјутерске графике анализирати садржину слике и применити на већ познате методе препознавања облика слова (на пример применом OCR). Захтев је израда једне класе Distiller-а која прихвата HTML страницу на којој се налази слика и која издваја слику. Израда класе Processor која има у себи уграђене механизме за одлучивање о садржају слова у слици. Класе које треба реализовати су Segmentation, Filter, Rotation, и треба користити пакете JOCR и GOCR. Користећи више расположивих метода класа Processor доноси одлуку о словима у оквиру слике користећи верзију edit distance алгоритма која је модификована у смислу коришћења меког одлучивања (soft decision) о вредности појединих слова у слици. Класа Segmentation омогућава изоловање појединих слова. Класа Rotation омогућава ротацију слова. JOCR и GOCR пакети служе за примену развијених метода детекције слова. Класа Filter има више намена које су одређене потребама конкретне имплементације са обзиром на тип CAPTCHA који се разматра. Обзиром да је GOCR писан у програмком језику C успешно писање имплементације подразумева писање native метода у Java-и која ће користити писани програм у C-у. Закључак рада је степен поузданости који се може постићи применом неког изабраног engine-а за конструкцију словних CAPTCHA.

Литература

1. Документација и source пројекта GOCR на адреси <http://jocr.sourceforge.net/index.html>
2. Документација и source пројекта GOCR на адреси <http://sourceforge.net/projects/javaocr/>
3. S. Mori, H. Nishida, H. Yamada, *Optical Character Recognition*, John Wiley & Sons, 1999.