

SubtitleNET

(dokumentácia k projektu na predmete VPPJ)

Vypracoval: Peter Dulačka
Akademický rok: 2010/2011, LS

1 Špecifikácia funkcionality:

Program obsahuje nasledujúcu funkcionality:

- Vyhľadávanie titulkov na základe vybraného video súboru
- Priame sťahovanie titulkov do zložky s videom

Program neobsahuje

- Možnosť priamo stiahnuť video/film
- Offline funkčnosť: žiadna interná databáza, všetko bude získavané z webu

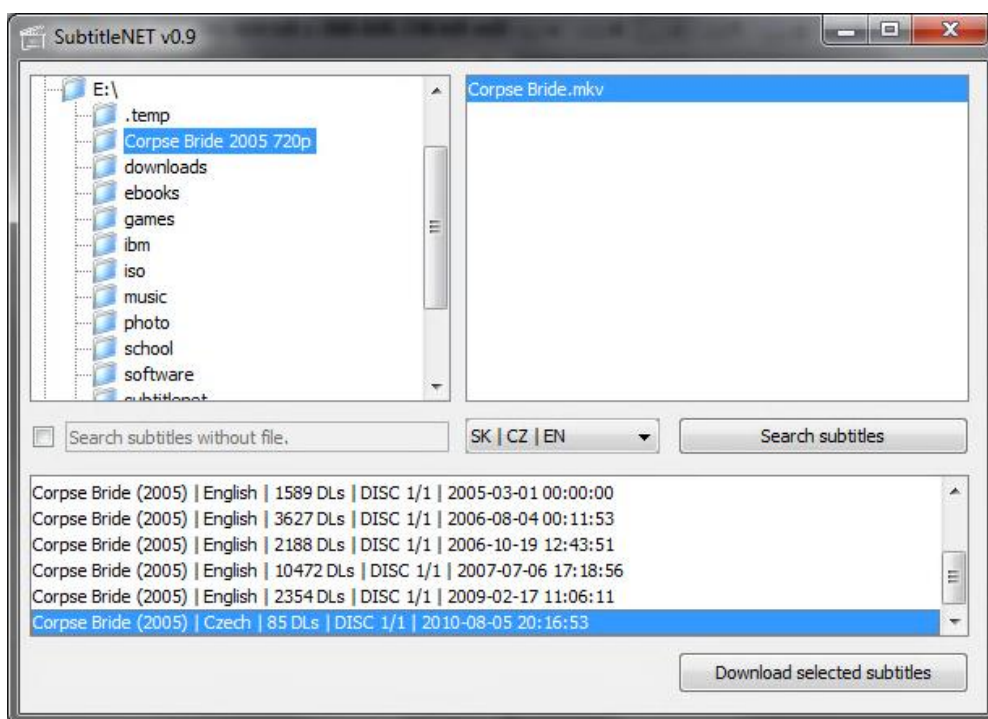
Žiadne zmeny oproti pôvodnej špecifikácii nenastali.

Technológie využité v programe:

- input/output
- swing
- xml dom
- http
- logging
- collections
- concurrency
- jar

2 Popis inštalácie a externé knižnice

Pre spustenie aplikácie stačí spustiť priložený .jar súbor (štandardne subtitled.jar). Pre správne fungovanie a lepšie identifikovanie príp. chýb je ale odporúčané spúšťať aplikáciu pomocou súboru run.bat, ktorý umožní logovanie činnosti. Spustený program je zobrazený na obr. 1.



obr. 1 Hlavná obrazovka SubtitleNET v0.9

Pri implementácii bola použitá externá trieda a to OpenSubtitlesHasher.java, ktorá bola prevzatá z adresy <http://trac.opensubtitles.org/projects/opensubtitles/wiki/HashSourceCodes>. Trieda slúži na výpočet hashu súbor, ktorý sa potom porovnáva s hashmi uloženými na serveroch OpenSubtitles.org. Druhá externá knižnica je org.apache.commons.lang, obsahujúca metódu na escapovanie stringov (kvôli korektnému využitiu xml). Knižnica je voľne dostupná na stiahnutie na adrese <http://www.java2s.com/Code/JarDownload/commons-lang-2.4.jar.zip>.

3 Implementačná dokumentácia

Jednotlivé metódy a triedy sú stručne opísané aj v javaDocu. Na obr. 2 je zobrazený diagram tried programu, ktorý obsahuje len najdôležitejšie metódy a atribúty (potrebné na pochopenie funkcionality).

pd.fiit.gui

GUI: Samotné grafické rozhranie, inicializácia komponentov a nastavenie listenerov. V tejto triede sú volania na handlers bežiacie vo vlastných threadoch.

TreeHandler: Zistí adresáre a súbory obsiahnuté vo vybranom adresári (pomocou FolderTraversara) a aktualizuje ako strom v ľavej časti GUI, tak list v pravej časti. Handler filtruje nájdené súbory len na video súbory (tie sú získané z dokumentácie OpenSubtitles.org API).

FolderTraversar: Nerekurzívne prechádza vybraný adresár a vracia zoznam súborov a adresárov.

WorkingDialog: Okno, ktoré sa zobrazuje počas komunikácie so serverom. Používaný v searchHandleri a downloadHandleri.

pd.fiit.subtitlenet

DownloadHandler: Dostane objekt „Subtitle“ v ktorom sa nachádzajú informácie potrebné na stiahnutie. Stiahne gzip archív do adresára k videu, archív rozbalí a odstráni. Výsledný súbor sa volá rovnako ako názov filmu.

LanguageHandler: Spravuje jazyky podporované API a konvertuje používateľský výber na reťazec potrebný pre API.

LoginHandler: Hneď po spustení programu sa snaží odoslať login request na server a získať token, ktorý je potrebný pre využívanie ostatných metód API. Po získaní odpovede zo servera využije „ResponseHandler“ na vytiahnutie hodnoty premennej „token“, ktorý je držaný ako atribút GUI.

Main: Nastavuje vzhľad programu (osobitný a zároveň štandardný pre každý OS), vytvára okno GUI a drží hodnotu aktuálnej verzie programu.

MethodWrapper: Využíva metódy triedy XmlRpc na vytvorenie štruktúr požadovaných xml-rpc a ich naplnenie podľa manuálu API.

OpenSubtitlesHasher: Externá trieda počítajúca hash video súboru.

ResponseHandler: Vytvorí z odpovede servera XML a v ňom nájde hodnotu požadovaných premenných (všetko podľa xml-rpc štandardu).

SearchHandler: Rozlišuje, ktorý spôsob vyhľadávania sa použije (hash, text), odosiela requesty na API server a vyťahuje potrebné údaje z odpovede servera do poľa typu „Subtitle“, ktoré potom vkladá do spodného listu a vracia ako výsledok threadu. Vyhľadávanie pomocou textu prebieha tak, že zadaný reťazec sa cez API server posunie na imdb a pre prvý nájdený výsledok sa potom hľadajú titulky. Vyhľadávanie pomocou hashu prebieha vypočítaním hashu videa, odoslaním requestu na API server, ktorý vráti priamo informácie o titulkoch sediace na daný hash.

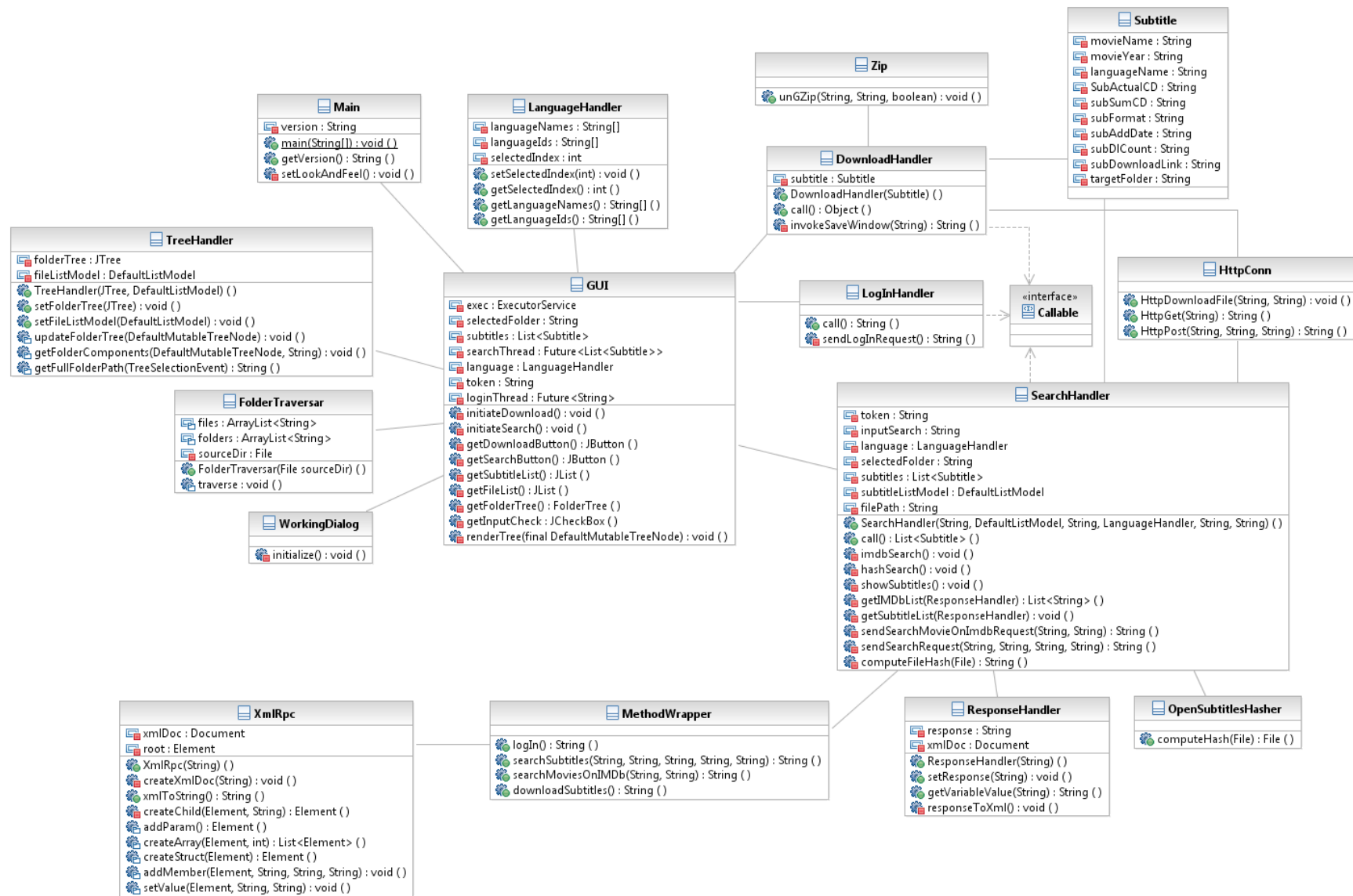
Subtitle: Štruktúra na ukladanie informácií o titulkoch.

XmlRpc: Trieda na jednoduché vytváranie xml-rpc štruktúr. Podrobne o tom ako štruktúry vyzerajú v špecifikácii xml-rpc (<http://www.xmlrpc.com/spec>).

pd.fiit.reusable

HttpConn: Trieda na odosielanie GET a POST requestov na server a sťahovanie súborov na disk. Naprogramovaná už dávnejšie, čo vysvetľuje horšiu kvalitu kódu. POST metóda nevyužíva knižnice javy, ale pracuje priamo so socketom a vytvára celú http hlavičku, kvôli minulým problémom s javovskou knižnicou. GET metóda v prípade vyhodenia výnimky vyskúša žiadosť ešte raz a skúsi, či nebola gzipovaná.

Zip: Trieda na prácu s archívami, momentálne obsahuje len metódu na rozbalenie .gz archívu.



obr. 2 Diagram tried

4 Prevádzková dokumentácia

Zoznam logovaných informácií so stupňom dôležitosti:

- prihlásenie na API server (INFO)
- nedostupnosť API servera, aj GET aj POST requesty (SEVERE)
- nemožnosť vypočítať hash vybratého súboru (SEVERE)
- snaha pripojiť sa na zlú URL (WARNING, štandardne nemá ako nastať)
- nemožnosť vytvoriť súbor s titulkami (SEVERE)
- nemožnosť vytvoriť prázdny xml-rpc dokument (SEVERE, štandardne nemá ako nastať)
- nemožnosť skonvertovať XML dokument do stringu (WARNING)
- úspešné stiahnutie titulkov (INFO)
- prerušenie login a search threadu (SEVERE)
- nemožnosť získať výsledok login alebo searchThreadu (WARNING)
- nemožnosť konverzie odpovede servera do XML (SEVERE)
- nemožnosť rozbalenia gzip archívu s titulkami (WARNING)

Samotné logovanie je možné nastaviť v súbore config.cfg, kde sa dá určiť aké levely udalostí sa budú logovať prípadne kam sa bude ukladať súbor s logmi. Súbor je pribalený v archíve s programom. V prípade neprítomnosti súboru sa nič neloguje.

Známe chyby a problémy

- program dokáže vyhľadávať len samostatné titulky, teda ak je film rozdelený na 2CD, treba vyhľadávať titulky osobitne pre každý video súbor; rovnako aj pri textovom vyhľadávaní je potrebné osobitne stiahnuť všetky titulky
- pri textovom vyhľadávaní a veľkom počte výsledkov trvá vyhľadávanie príliš dlho, odporúča sa nepoužívať tento typ vyhľadávania najmä pri seriáloch a používať radšej vyhľadávanie podľa súboru na disku
- program je závislý na vzdialenom API serveri, ktorého dostupnosť je neovplyvniteľná; v prípade oznamu, že sa nedá pripojiť na server sa odporúča skontrolovať dostupnosť servera OpenSubtitles.org, keďže nie je ideálna a server máva výpadky (a to hlavne počas „peak hours“ zhruba od deviatej večer do polnoci a hlavne počas víkendov).
- WorkingDialog má nulovú výšku kvôli problémom s vykresľovaním samotného dialogu. Toto je cieľené keďže súčasná implementácia nedokáže vykresliť obsah okna počas pripájania sa na server a po pripojení sa dialog zatvára.