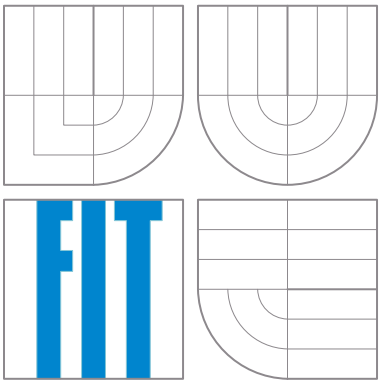

IRC bot s logováním přes SYSLOG protokol

Dokumentace k projektu do předmětu ISA

16. listopadu 2014

Autor: Vojtěch Havlena (xhavle03)
Fakulta Informačních Technologií
Vysoké Učení Technické v Brně

Obsah

1 Úvod do problematiky

Cílem projektu do předmětu „Síťové aplikace a správa sítí“ je implementace jednoduchého IRC bota, který umožňuje logovat vybrané zprávy na daný server. Výběr příchozích zpráv pro logování je podmíněn přítomností alespoň jednoho slova z daného seznamu. Pokud seznam slov není zadán, logují se všechny přijaté zprávy.

1.1 Komunikace s IRC serverem

IRC je otevřený protokol [1], který umožňuje komunikaci v reálném čase po Internetu. Komunikace s IRC serverem probíhá pomocí TCP přes port zadaný uživatelem. Výchozím portem je však 6667. Po navázání spojení s IRC serverem je zaslána následující dvojice zpráv, sloužící pro identifikaci a přihlášení uživatele (pouze pro servery bez registrace a hesla):

```
NICK <nick>
USER <username> <hostname> <servername> <realname>
```

V našem případě všechny položky odpovídají loginu, tedy `xhavle03`. Po úspěšném přihlášení pomocí příkazu `JOIN <channel>`, dojde k připojení uživatele do kanálu, který je zadán jako parametr programu. Během činnosti ircbota, IRC server pravidelně generuje zprávy `PING`, na které musí ircbot co nejrychleji odpovědět zprávou `PONG`. Tyto zprávy slouží ke zjištění serveru, zda je klient stále připojen. Zprávy, které se mají potencionálně logovat přicházejí příkazem `PRIVMSG` a `NOTICE`. Z celé zprávy serveru je nutné zjistit nick odeílatele a obsah samotné zprávy. Ukončení komunikace se IRC serveru oznamuje zprávou `QUIT`.

1.2 Logování zpráv

Logování probíhá pomocí UDP přes port 514. Pro komunikaci se využívá protokolu SYSLOG [2]. Tento protokol je používán pro přenos upozornění na údálosti po síti. Formát zprávy posílaného na logovací server je dán následujícím předpisem:

```
<134>Mmm dd hh:mm:ss <IP adresa odesílatele> ircbot [<<nick>>]:
                                <Text zprávy>
```

Kde číslo 134 zohledňuje požadované hodnoty facility a severity. Další částí je časová známka, kde `Mmm` je anglická zkratka aktuálního měsíce, `dd` je číslo dne (v případě, že je menší než 10, je nutné je reprezentovat mezerou a číslem) a zbytek je hodnota aktuálního času. V případě, že je logována chybová zpráva, je hodnota `<nick>` vynechána. Pokud je zachycena chybová zpráva z IRC serveru je celý text, který popisuje vzniklou chybu zalogován.

2 Návrh aplikace

Aplikace je logicky členěna do několika tříd, které zapouzdřují operace pro komunikaci s IRC a logovacím serverem. Samotný hlavní program se skládá z řady podpůrných funkcí, které zpracovávají parametry spuštění, parsují chybové hlášky, starají se o korektní připojení (odpojení) od serverů apod.

Pro uložení parametrů spuštění se využívá struktura `PARAMS`. Mimo IRC, SYSLOG server, port a kanál pro připojení, obsahuje struktura také seznam klíčových slov. V případě, že nejsou zadána žádná klíčová slova, jsou logovány všechny příchozí zprávy.

2.1 Třída IrcParser

První třídou je třída `IrcParser`, která parsuje zprávy posílané IRC serverem. Parsování probíhá pomocí gramatiky, která specifikuje formát zprávy (viz. dále). Třída umožňuje rozlišování prefixu, parametrů a samotného obsahu zprávy. Toto je nezbytně nutné pro pozdější zalogování odesílatele a obsahu zprávy (přezdívka odesílatele je uložena právě v prefixu zprávy). Třída obsahuje také podpůrnou metodu `GetNickFromPrefix` pro parsování přezdívky z prefixu zprávy.

2.2 Třída SyslogHandler

Druhou vytvořenou třídou je `SyslogHandler`. Tato třída zajišťuje kompletní obsluhu komunikace s logovacím serverem, což zahrnuje obsluhu připojení k serveru, generování logovacích zpráv (metoda `GenerateMessage`) a jejich odeslání přes UDP (metoda `SendLog`). Tyto operace jsou implementovány v členských metodách. Pro vygenerování validní logovací zprávy je nutné zjistit IP adresu odesílatele, za kterou se považuje IP adresa prvního nalezeného síťového rozhraní (mimo loopback). Toto zajišťuje metoda `GetHost`.

3 Popis implementace

Hlavní funkce programu je, mimo nezbytné inicializace proměnných zodpovědných za síťovou komunikaci, implementována jako potenciálně nekonečný cyklus. Na začátku každé iterace dojde k blokujícímu čtení příchozí zprávy (právě jeden řádek) od IRC serveru. Zpráva je posléze zpracována třídou `IrcParser`. V případě, že byla zaslána zpráva `PRIVMSG` nebo `NOTICE`, zkontroluje se obsah zprávy na přítomnost vyhledávaných slov a případně se zpráva zaloguje. Vyhledává se podle shody na podřetězec. Zmiňovanou operaci zalogování zajišťuje metoda `SyslogHandler::SendLog`. V případě, že je zachycena chybová zpráva z IRC serveru (číslo 400 a výše) nebo chyba při libovolné síťové operaci, je program ukončen.

Aby se program korektně odhlásil od IRC serveru, je implementována obsluha signálů SIGINT a SIGTERM (Ctrl+C), po jehož zachycení dojde ke korektnímu ukončení programu.

3.1 Parsování IRC zpráv

Jak již bylo zjíněno dříve, parsování probíhá ve třídě `IrcParser` a je založeno na gramatice specifikované v RFC 1459 [1]. Zpracovávání probíhá po jednotlivých řádcích. Parsování probíhá v několika krocích:

1. Zjištění prefixu zprávy a následné rozpoznání případného nicku odesílatele. Nick je část prefixu po znak '!' nebo '@', eventuálně je možné, aby nick obsadil celý prefix.

2. Parsování příkazu (popř. trojciferného čísla chyby).

3. Parsování parametrů zprávy (zbytek řetězce zprávy). Třída obsahuje i členskou metodu `GetMessageFromParams`, která umožňuje zjistit obsah zprávy z těchto parametrů. Obsah zprávy je posledním parametrem (parametry jsou odděleny mezerami) nebo je uvozen znakem ':'.

3.2 Odesílání logovacích zpráv

Před samotným odesláním logovací zprávy je nutné zprávu vygenerovat. Formát zprávy vyžaduje speciální formát data (viz. kapitola 1), který je generován v metodě `GenerateTimestamp` třídy `SyslogHandler`.

Logují se zprávy obsahující zadané klíčové slovo, ale i vzniklé chyby. V případě logování chybových zpráv se ve formátu zprávy vynechává políčko přezdívky (vyplněno pouze při logování zprávy od uživatele). Logované chyby jsou dvojího druhu: chyby z IRC serveru (např. snaha o použití již existující přezdívky) a chyby, které vznikly při síťové komunikaci (připojování k IRC serveru, zasílání zpráv apod).

4 Návod k použití

Program podporuje následující syntaxi spuštění:

```
ircbot <host>[:<port>] <channel> <syslog_server> [<list>]
```

`<host>` – název serveru
`<port>` – číslo portu, na kterém naslouchá server (výchozí hodnota je 6667)
`<channel>` – jméno kanálu, do kterého se ircbot připojuje
`<syslog_server>` – ip adresa logovacího serveru
`<list>` – seznam hledaných slov oddělených středníkem

4.1 Příklad spuštění

```
./ircbot irc.nomi.cz "#fit" 127.0.0.1 "lan;wan;tcp;ip;isa"
```

Příklady zpráv zalogovaných na serveru:

```
Oct 12 19:56:56 192.168.0.197 ircbot <gren258>: tcp je spojovana
transportni sluzba
Oct 15 17:55:15 192.168.0.197 ircbot: Nickname is already in use.
```

Reference

[1] Oikarinen J. a Reed D. *Internet Relay Chat Protocol (RFC 1459)*. Květen 1993. `https://tools.ietf.org/html/rfc1459`

[2] Lonvick C. a Cisco Systems, *The BSD syslog Protocol (RFC 3164)*. Srpen 2001. `http://www.ietf.org/rfc/rfc3164.txt`