

AKADEMIA GÓRNICZO-HUTNICZA

**Wydział Elektrotechniki, Automatyki,
Informatyki i Elektroniki**



Autoreferat Rozprawy Doktorskiej

**Analiza wyszukiwania i wymiany informacji
w złożonych systemach
e-learningowych**

mgr inż. Tomasz M. Orzechowski

Promotor:

Prof. dr hab. inż. Andrzej Dziech

Akademia Górniczo-Hutnicza

Recenzenci:

Prof. dr hab. inż. Zygmunt Wróbel

Uniwersytet Śląski

Prof. dr hab. inż. Andrzej Pach

Akademia Górniczo-Hutnicza

Kraków 2009

Wprowadzenie

Systemy e-learningowe zyskują coraz większą popularność w wielu obszarach edukacyjnych. Pozwalają bowiem zarówno kształcić na odległość (ang. *distance learning*), jak i pozwalają twórcom treści elektronicznych (ang. *e-Contents*) na tworzenie wyspecjalizowanych centrów szkoleniowych.

Wpływ e-learningu na kształcenie będzie się rokrocznie powiększał. Wynika to zarówno z samego rozwoju infrastruktury jak i sposobu oddziaływania tej formy przekazu. Herbert Marschall McLuhan, medioznawca żyjący w latach 1911-1980, podkreślał, że media są podstawowym czynnikiem wszelkich zmian społecznych. Jego teoria, iż „medium jest przekazem” wskazuje na istotny wpływ charakteru samego środka przekazu na przenoszoną treść oraz na jakość jej odbioru.

Systemy e-learningowe potrafią wykorzystać opisywany wpływ sposobu przekazywania treści angażując niemal całkowicie użytkownika, który będąc w ciągłej interakcji z otrzymywaną treścią jest w stanie szybciej i skuteczniej przyswoić sobie zawartą w niej wiedzę [39]. Prawidłowość tę potwierdzają badania przeprowadzone przez IBM [14] które wskazały na dwie podstawowe korzyści: efektywność szkoleń oraz na wyraźne obniżenie kosztów szkoleń. Przykładowo zastosowanie technik e-learningowych pozwoliło zmniejszyć tej firmie koszty szkoleń o 50-70% i skrócić czas szkolenia z 11 do 9 tygodni przy równoczesnym wzroście produktywności tak wykształconych pracowników o 100-170%. Nic więc dziwnego, że techniki e-learningowe stają się popularne i są wspierane na całym świecie.

Promowanie technik e-learningowych ma miejsce w krajach Unii Europejskiej m.in. w ramach Funduszy Strukturalnych UE (lata: 2007–2013), czy w ramach programu „Uczenie się przez całe życie” – „**Lifelong Learning Programme (LLP)**” [33].

Podstawową jednostką treści elektronicznej jest obiekt wiedzy (ang. *LO – Learning Object*) którym może być zarówno treść elementarna (ang. *Learning Asset*), przykładowo: pojedyncze zdjęcie, film, plik tekstowy, czy obiekt złożony, składający się z wielu LO odpowiednio ze sobą powiązanych, a stanowiących np. pojedynczą jednostkę lekcyjną, czy cały kurs. Dobrze opisany LO w oparciu o standard IEEE LOM (ang. *Learning Objects Metadata*) [40] może być współdzielony nie tylko przez użytkowników danej platformy e-learningowej lecz mogą go wykorzystywać użytkownicy różnych innych platform. Ponadto część obiektów może być dostarczana przez firmy ukierunkowanie wręcz tylko na tworzenie samych zasobów edukacyjnych, które są udostępniane na zasadach licencji użytkownikom zewnętrznych platform e-learningowych.

Podjęcie tej tematyki jest szczególnie ważne z uwagi na znaczącą rolę systemów e-learningowych w procesie zdobywania wiedzy. Intensywny wzrost objętości informacji przekazywanej i zawartej w tego typu systemach powoduje, iż zachodzi konieczność opracowania modelu systemu wyszukiwania, pozwalającego na jego łatwą rozbudowę, a wyposażonego w szybkie, wydajne metody zapewniające użytkownikom skuteczne dotarcie do poszukiwanej informacji.

Tematyka rozprawy obejmuje zagadnienia dotyczące systemów wyszukiwania informacji ze szczególnym uwzględnieniem systemów wykorzystujących platformę agentową i stanowiących fragment złożonego rozproszonego systemu e-learningowego.

Niniejsza rozprawa ma również na celu szczegółowe omówienie aspektów związanych ze skutecznym i szybkim wyszukianiem informacji w systemach e-learningowych. Jednym z celów pracy jest także wskazanie praktycznych obszarów zastosowań systemów wieloagentowych w rozproszonych systemach wyszukiwania na przykładzie realizowanych przez autora prac badawczych.

Tezy rozprawy

W niniejszej pracy postawiono dwie tezy:

1. Zastosowanie podejścia spersonalizowanego zapewnia skuteczny dostęp do żądanej informacji spełniającej oczekiwania konkretnego użytkownika w złożonych wielojęzycznych systemach e-learningowych.
2. Wprowadzenie systemów wieloagentowych zapewnia łatwiejsze zarządzanie złożonym systemem wyszukiwania wielojęzycznej informacji.

W celu udowodnienia pierwszej tezy, w pracy zaproponowano wykorzystanie specyfiki systemów e-learningowych w tym zestandaryzowanego opisu obiektów wiedzy oraz wprowadzono opis profilu użytkownika komplementarnego wobec opisu obiektów. Następnie zaproponowano możliwości uwzględniania danych opisujących użytkownika dla podniesienia jakości wyszukiwania. W tym celu opracowano kilka niezależnych scenariuszy analizy informacji opisujących użytkowników (profil, zachowanie) w aspekcie opisu samych obiektów wiedzy.

Tego typu podejście do wyszukiwania spersonalizowanego w praktyce nie było oferowane przez żaden istniejący system e-learningowy i ma, jak się wydaje, charakter nowatorski.

Istotnym problemem jest znalezienie skutecznego rozwiązania zagadnienia wzrostu objętości zasobów informacyjnych w wielojęzycznych systemach e-learningowych. Druga teza stanowi propozycję wprowadzenia systemu wieloagentowego w celu wydajnego rozdziału realizowanych zadań przez system gromadzący dane i asystujący użytkownikowi w procesie wyszukiwania. System gromadzący dane może przetwarzać zasoby po stronie lokalnego repozytorium przy zastosowaniu agentów mobilnych. Systemy wieloagentowe nie były do tej pory stosowane w praktyce w systemach e-learningowych. Zaproponowane nowatorskie rozwiązanie gwarantuje łatwą rozbudowę systemu oferując nie tylko możliwość rozpraszania komponentów platformy agentowej w sieci, ale również zapewnia rozbudowę funkcjonalności proponowanej struktury w przyszłości.

1. Wymiana informacji w rozproszonych systemach e-learningowych.

Popularność systemów e-learningowych staje się coraz większa. Systemy te wypełniają bowiem zarówno braki w dostępie do standardowych metod kształcenia, jak i zapewniają wydajniejszy, pogładowy, przekaz. Nowoczesne systemy e-learningowe potrafią ponadto przeanalizować umiejętności osoby kształczonej i tak dopasować tempo pracy aby utrwalenie wiedzy dokonało się możliwie jak najszybciej.

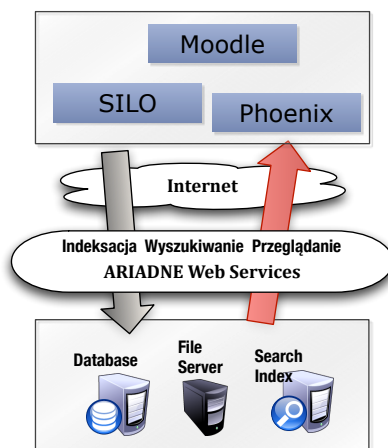
Stały wzrost liczby obiektów dydaktycznych powoduje, że w celu zapewnienia wysokiej jakości kształcenia, kluczowe staje się udostępnianie ich w globalnej sieci repozytoriów LOs oraz repozytoriów LOMs.

Potrzeba zapewnienia wymiany danych edukacyjnych jest znana od ostatnich lat. Do najbardziej znanych inicjatyw można zaliczyć:

- **ARIADNE Foundation** w Europie (The European Association for Knowledge Sharing and Reuse) www.ariadne-eu.org
- **Multimedia Educational Resources for Learning and Online Teaching** (MERLOT) w USA, www.merlot.org
- **LORNET** w Kanadzie, www.lornet.org
- **education.au** w Australii, www.educationau.edu.au/jahia
- **National Institute of Multimedia Education** (NIME) w Japonii, www.nime.ac.jp

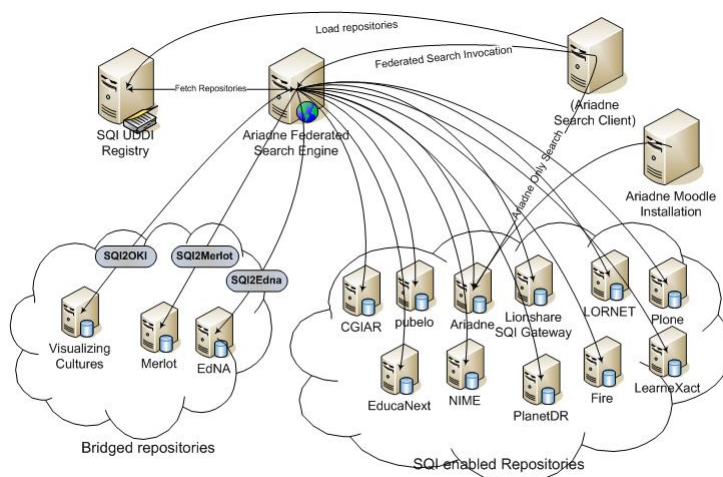
Organizacje te obecnie utworzyły wspólną grupę **The Global Learning Objects Brokered Exchange** (GLOBE) mającą na celu integrować wszelkie prace standaryzacyjne zapewniające tworzenie i rozwijanie otwartych standardów wymiany danych edukacyjnych na jednej platformie brokerskiej.

Sieć Ariadne stanowi przykład rozproszonej sieci repozytoriów w UE. Podstawowym elementem struktury jest repozytorium które może być przeszukiwane w oparciu o SQL. Trójwarstwowa struktura powoduje, że z punktu widzenia użytkownika i jego aplikacji sieć Ariadne zostaje ukryta pod warstwą pośrednią (Rys. 1) [42].



Rysunek 1. Warstwa pośrednia w komunikacji klient-repozytorium

Wprowadzony model architektury pozwala na tworzenie Federacji Repozytoriów (Rys. 2) [42], które mogą być przeszukiwane równocześnie w oparciu o wyszukiwanie federacyjne (ang. *Federated Search*) [43]. Wyszukiwanie Federacyjne w ujęciu asynchronicznym polega na odebraniu wyników od klienta przez Brokera, rozesłanie przez Brokera zapytań do zarejestrowanych repozytoriów, odebranie poszczególnych wyników przychodzących z repozytoriów i systematyczne wysyłanie ich do klienta [45].



Rysunek 2. Model Sieci Ariadne

Dążenie do ujednoczenia powstałych niezależnie od siebie standardów doprowadziło do ujednoczania standardów wprowadzanych historycznie przez różne organizacje i tak:

- **W zakresie opisu obiektów wiedzy.**

Konieczność przeszukiwania zasobów w celu odnalezienia poszukiwanych obiektów wiedzy przy coraz większym ich przyroście oznacza, iż różne systemy winny stosować ten sam standard opisu LO [2, 9]. W chwili obecnej standard IEEE LOM jest powszechnie uznany. Został też przyjęty zarówno przez IMS jak i SCORM. Specyfikacje rozwijane przez te organizacje wskazują na proces migracji do IEEE LOM 1484.12 (IMS Learning Resource Meta-data Specification v.1.3). Warto zwrócić uwagę, że SCORM Metadata standards powstał na bazie standardu IEEE 1484.12.

- **W zakresie opisanie struktury treści kursów.**

Istnieją dwa szeroko przyjęte standardy pakowania danych, standard AICC oraz IMS GC. Wersja 1.2 specyfikacji SCORM zawiera zaadoptowany standard IMS GC. Standard ten jest powszechnie stosowany i zalecany.

- **W zakresie komunikacji studenta z systemem e-learningowym.**

Standardy komunikacji określają sposoby uruchamiania kursów oraz wymiany komunikatów trakcie nauki dotyczących jej przebiegu. Istniały dwa główne standardy:

- **AICC** – *Guidelines and Recommendations (AGRO06 i AGRO10)*,
- **SCORM** – *Runtime Environment (RTE)*.

Standard SCORM jest on zalecany do stosowania.

- **W zakresie dostępu do opisu obiektów wiedzy.**

W chwili obecnej, w zależności od potrzeb, wykorzystuje się standard SQI lub OAI.

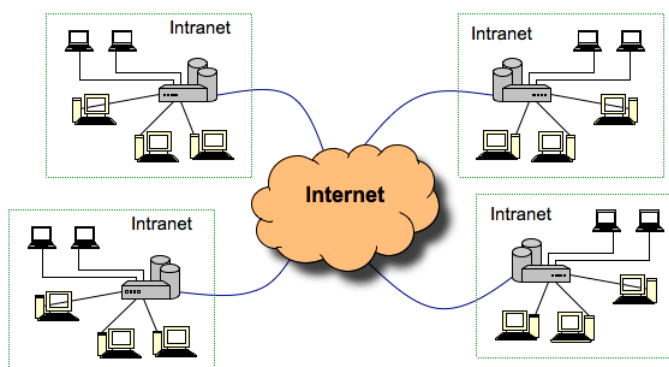
Tworzenie rozproszonych systemów e-learningowych oraz wsparcia dla tworzenia i wymiany obiektów dydaktycznych wiąże się z koniecznością wprowadzenia technik oznaczania jakości materiałów oraz wprowadzenia szybkich i wydajnych mechanizmów wyszukiwania obiektów. Wyszukiwanie powinno uwzględniać charakterystyki osoby szukającej, jak i jakości samego obiektu wiedzy.

Rozprawa stanowi propozycję takiego systemu, zaś szczegóły implementacji zostały przedstawione w rozdziale: „Agentowy System Wyszukiwania”, str. 16.

2. Systemy wieloagentowe

Twórcy systemów informatycznych dość często stają przed pytaniem o sposób organizacji zasobów i dostępu do nich. Każde z istniejących rozwiązań tj. podejście scentralizowane czy zdecentralizowane posiada zarówno swoje wady jak i zalety. Model scentralizowany pozwala na łatwiejsze zarządzanie zasobami charakteryzującymi się niewielką zmiennością, zaś model zdecentralizowany pozwala na łatwość dostępu do danych które mogą być rozproszone w sieci. Czasem uważa się, że być może jedyną drogą do realizacji systemów zdecentralizowanych są systemy agentowe [6]. Liczne analizy wskazują jednak na problemy realizacji systemów zdecentralizowanych w oparciu o systemy agentowe i agentów mobilnych [6, 8, 48].

Autor rozprawy zwraca uwagę na możliwość wprowadzenia agentów mobilnych pozwalających przeszukiwać rozproszone zasoby sieci widziane jako zbiór scentralizowanych zasobów lokalnych danych, zgromadzonych w repozytoriach podłączonych do Internetu (Rys. 3)[19].



Rysunek 3. Szkielet systemu wyszukiwania danych

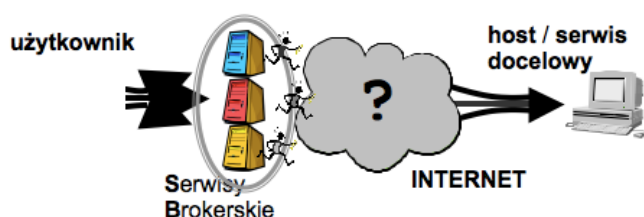
Przedstawiony system ma charakter hybrydowy. Układ ten odpowiada lokalnym strukturom repozytoriów wiedzy (zawierających obiekty wiedzy LOs i ich metadane LOMs). Zaproponowany w 2001 roku model dobrze wpisuje się w tendencje tworzenia globalnego systemu współdzielenia danych edukacyjnych zgodnie z obecnym trendem rozwoju platform e-learningowych.

Wątpliwości dotyczące wydajności takich systemów przedstawiono szerzej w rozprawie, a wyniki są zgodne z badaniami autora niniejszej rozprawy, który wykazał, że korzyści ze stosowania technik agentowych pojawiają się dopiero przy przekroczeniu określonej liczby zadań (zapytań) wykonywanych przez agenta mobilnego oraz określonej liczby rozproszonych zasobów, które są przez agenta przeszukiwane [22].

Możliwość zastosowania technik negocjacji [6, 8] jako naturalnej cechy agentów autonomicznych, do obsługi realizacji usług QoS w sieciach komputerowych została również zaproponowana przez autora niniejszej rozprawy [26, 27].

Wydaje się, że technologia agentów mobilnych stanowi dobre rozwiązanie do tworzenia serwisów brokerskich w sieciach opartych o zasady rynkowe (rozumiane tu jako sieci w których dynamicznie pobierane są opłaty za korzystanie z usług sieciowych). Zaproponowane serwisy mogą gromadzić i przetwarzać informacje o stanie istniejących połączeń, ich charakterystykach, kosztach korzystania, itp., a w konsekwencji dostarczać użytkownikowi końcowemu konkretną ofertę przesyłu danych, czy dostępu do określo-

nych zasobów z uwzględnieniem zadanych kosztów, czy parametrów QoS [28].
Serwis brokerski stanowi serwis pośredniczący, zlokalizowany pomiędzy użytkownikiem a serwisem docelowym (Rys. 4).



Rysunek 4. Lokalizacja serwisów brokerskich

Serwis taki zapewnia użytkownikowi kompleksową obsługę dostępu do danych zasobów poprzez zaprezentowanie użytkownikowi konkretnej oferty i po jej opłaceniu zrealizowaniu usługi.

Funkcjonalność proponowanego serwisu dokonuje się dwuetapowo:

- Pierwszy etap polega na gromadzeniu informacji i jej przetwarzaniu,
- Drugi zaś sprowadza się do obsługi klienta.

Zbieranie informacji o strukturze sieci, oferowanych usługach i możliwościach rezerwacji połączeń przy zagwarantowanych parametrach połączeń ma charakter statyczny. W przypadku agentów mobilnych istnieje możliwość wprowadzenia dynamiki. Dynamika ta polega na wysyłaniu agentów, którzy kontaktując się z agentami statycznymi danego providera mogą negocjować koszty korzystania z danych usług. Co więcej sam provider ma możliwość (przy przyjęciu aukcyjnego modelu) sprzedaży serwisom brokerskim usług za najlepszą wylicytowaną kwotę.

Systemy wieloagentowe mogą także być stosowane do prowadzenia symulacji. Dany agent w takim systemie może być wyposażony w określone cechy, może podejmować decyzje i ewoluować. Symulacje mogą dotyczyć między innymi procesów rynkowych. Kompleksowe omówienie technologii agentowej można znaleźć m.in. w [6, 8].

Zainteresowania badawcze autora niniejszej rozprawy nie są jednak ukierunkowane na obszar symulacji, lecz na wykorzystanie systemów wieloagentowych do tworzenia aplikacji rozproszonych.

Tworzenie aplikacji rozproszonej zorientowanej na agentów wymaga doboru odpowiednich narzędzi AOL (ang. *Agent-Oriented Languages*). Istnieje duża liczba platform realizujących paradygmat programowania zorientowanego na agentów.

Model zarządzania wyspecyfikowany przez FIPA (Management Reference Model), stanowiący rdzeń platform zgodnych ze specyfikacją FIPA i gwarantuje poprawność zarządzania agentami w tym: transferu kodu, jego obsługi lokalnej i zdalnej, a także poprawnej komunikacji pomiędzy agentami oraz ich otoczeniem.

Autor rozprawy przeprowadził szczegółową analizę wybranych platform agentowych, w szczególności: Agent Development Kit (www.tryllian.com), April Agent Platform, Comtec Agent Platform, FIPA-OS, Grasshopper, JADE, JAS (Java Agent Services API) oraz ZEUS. Przy analizie platform agentowych wspierających standard FIPA zwrócono także uwagę na fakt, czy jest ona w dalszym ciągu rozwijana, oraz czy jest oferowana bezpłatnie dla celów edukacyjnych.

Platforma JADE (Java Agent DEvelopment Framework) [4, 15] została ostatecznie wybrana jako narzędzie do realizacji testowego modelu agentowego systemu wyszukiwania spersonalizowanego opisywanego w rozdziale 3 rozprawy. Platforma ta wspiera specyfikacje FIPA, a także jest ciągle rozwijana. Należy podkreślić, że jest ona również bardzo popularna w środowiskach naukowo-badawczych z uwagi na wysoką wydajność działania [7, 37].

3. Problemy inteligentnego wyszukiwania informacji

Rozwój globalnej sieci komputerowej, łatwość umieszczania i współdzielenia w niej informacji powoduje wzrost potrzeby na wprowadzanie coraz bardziej skutecznych metod wyszukiwania informacji. Dziś obserwujemy też rozszerzanie obszarów, których dotyczy proces wyszukiwania. Obecnie obejmuje on nie tylko tekst ale również analizę treści multimedialnych.

W niniejszej pracy skupiono się na systemach wyszukiwania informacji bazującym wprawdzie na danych tekstowych, lecz zamierzeniem autora jest analiza metadanych stanowiących w istocie zestandaryzowany opis dowolnych danych interesujących z punktu widzenia osoby szukającej.

Skuteczność działania systemu charakteryzuje przede wszystkim jakość dostarczonych użytkownikowi wyników wyszukania w stosunku do zapytania użytkownika, a w drugiej kolejności czas potrzebny na przedstawienie użytkownikowi tych wyników. W związku z tym twórcy systemów wyszukiwania przede wszystkim koncentrują się na podniesieniu jakości działania algorytmów poprzez powszechnie stosowanie metod indeksacji dokumentów oraz wprowadzeniu reprezentacji zarówno samych dokumentów (i ich opisów) jak i zapytania użytkownika w postaci wektorów w przestrzeni n -wielowymiarowej (ang. *vector space model*) (1) dla określenia zgodności zapytania i treści dokumentu [34-36]:

$$\vec{v}_m = (w_{1,m}, w_{2,m}, \dots, w_{i,m}, \dots, w_{n,m}) \quad (1)$$

gdzie: $w_{i,m}$ stanowi wagę danego terminu i w dokumencie m ,

n stanowi liczbę terminów indeksowych reprezentujących dany dokument.

Waga danego terminu w modelu uproszczonym opartym o częstości występowania terminów indeksowych (ang. *Term Count Model*) jest równa tej częstości [34, 35]:

$$w_{i,m} = \varphi_{i,m} \quad (2)$$

gdzie: $\varphi_{i,m}$ stanowi częstość występowania danego terminu i w dokumencie m ,

W przypadku grupy dokumentów i wspólnej listy terminów indeksowych przyjęto, że jeśli dany termin indeksowy nie występuje w dokumencie, waga terminu wynosi zero.

Coraz częściej obecnie stosuje się także identyczną formę dla reprezentacji opisu zainteresowań użytkownika, co pozwala na lepsze dopasowanie („zrozumienie”) intencji pytającego [17, 30].

Warto również zwrócić uwagę na potrzeby tworzenia architektury zapewniającej dostęp użytkownikom do rozproszonych danych poprzez stworzenie spójnego rozproszonego systemu wyszukiwania informacji. Propozycja takiego systemu została zaproponowana przez autora rozprawy [19, 21, 23, 25].

Poza zastosowanym przez autora podejściem w systemach stosowane są też następujące modele wyszukiwania informacji: wektorowy oraz probabilistyczny [47]. Modele te uwzględniają stopień dopasowania pomiędzy zapytaniem, a wektorową reprezentacją dokumentu opisywaną wcześniej jako wynik dostarczają użytkownikowi listę dokumentów wraz z miarą dopasowania dokumentów do zapytania. Wartość dopasowania jest wyliczana poprzez porównanie wektorów dokumentów oraz wektora terminów zapytania, uwzględniając także rozkład wszystkich terminów indeksowych w całym zbiorze dokumentów. Rozwiązanie to nie zostało jednak uwzględnione z uwagi, iż przede wszystkim w pracy koncentrowano się na szybkości działania metod wyszukiwania. Analiza dotyczyła aspektów wyszukiwania spersonalizowanego, a nie na rozbudowy algorytmów wyszukiwania treści.

Metody wyszukiwania w rozproszonych systemach repozytoriów

W zależności od konstrukcji systemu rozproszonego autor rozprawy zaproponował następujące podejścia [19]:

- dla lokalnego wyszukiwania danych w systemie lokalnym stanowiącym część złożonego systemu rozproszonego.
- dla wyszukiwania danych w odległych repozytoriach systemu rozproszonego.

Oba podejścia zostały szczegółowo omówione w rozprawie.

Komunikacja z repozytoriami.

Dostęp do lokalnych repozytoriów można realizować w oparciu o SQL lecz wymaga to znajomości struktury przechowywanych danych w repozytorium co w praktyce unieumożliwia wydajne tworzenie architektury wyszukiwawczej. Problemy te nasilają się wraz ze wzrostem liczby repozytoriów udostępnianych szerszej liczbie użytkowników. Najczęściej spotykanymi standardami zapewniającymi szybki dostęp do repozytoriów są: SQI (ang. *Simple Query Interface*) [38] oraz OAI-PMH (ang. *The Open Archives Initiative Protocol for Metadata Harvesting*) [41].

Oba podejścia zostały uwzględnione przez autora rozprawy.

Systemy rekomendacyjne

Celem tworzenia systemów rekomendacyjnych jest podnoszenie jakości wyszukiwania zapewniającej dostarczanie konkretnemu użytkownikowi takich wyników wyszukiwania które będą najlepiej dopasowane do jego indywidualnych oczekiwań.

Systemy rekomendacyjne znajdują szerokie zastosowanie wszędzie tam gdzie tzw. inteligencja grupowa (ang. *Collective Intelligence – CI*) może zapewnić użytkownikom łatwiejsze i szybsze dotarcie do poszukiwanej przez nich informacji.

W systemach rekomendacyjnych zazwyczaj informacja jest przedstawiana w formie macierzy opisujących rekomendację bądź jej brak przez użytkowników dla każdego dokumentu/obiektu w kolekcji. Wartość określa siłę rekomendacji. Wyróżnia się dwa sposoby pobierania danych od użytkownika:

- jawny (explicit) – użytkownik ocenia dany dokument jawnie, lub
- niejawnie (implicit) – n.p. poprzez fakt pobrania dokumentu, zakupu produktu, itp..

Szczegółowa analiza istniejących systemów rekomendacyjnych jest opisana w [17, 32]. Autorzy tej pracy dokonali klasyfikacji systemów i wyróżnili następujące podstawowe ich rodzaje:

- filtrowanie demograficzne – DF (ang. *demographic filtering*);
- kolaboratywne – CF (ang. *collaborative filtering*);
- i na podstawie zawartości – CBF (ang. *content-based filtering*).

Warto wspomnieć, że czasem przedstawia się systemy rekomendacyjne jako podzbiór systemów filtrowania informacji (ang. *Information Filtering – IF*) [18].

Filtrowanie kolaboratywne określane też asocjacyjnym nie analizuje treści poszczególnych elementów wyszukiwanych czy wybieranych przez użytkownika lecz skupia się na wychwytywaniu podobieństw pomiędzy użytkownikami.

CF może być postrzegany jako specyficzne zastosowanie procesu eksploracji danych (ang. *data mining*), gdzie bazując na wcześniejszych rekomendacjach innych użytkowników przewiduje się które spośród zbioru obiektów elementy będą najbardziej interesujące dla danego użytkowników [13]. Użytkownikowi będą rekomendowane te obiekty, które są najbardziej popularne (informacja zwrotna) lub najwyżej oceniane przez inne osoby o podobnych zainteresowaniach. Szerokie spektrum dotyczące systemów CF opisano w [1, 17].

System zaproponowany przez autora niniejszej rozprawy gromadzi informacje o zainteresowaniach użytkowników w sposób jawny i niejawny. Sposób jawny polega na ocenianiu wyszukanych obiektów wiedzy przez poszczególnych użytkowników. Sposób niejawny na liczeniu liczby kliknięć obiektów na liście wyników (powodujące wyświetlenie strony ze szczegółami danego obiektu) oraz pobrania samych obiektów. Dominujące cechy macierzy ocen/zdarzeń ekstrahowane są przy pomocy metod statystycznych pozwalających na redukcję ilości wymiarów danych. Dane te następnie służą do grupowania użytkowników o podobnych preferencjach w klastry.

4. Wyszukiwanie informacji w systemach e-learningowych

Problemem wszystkich systemów opartych o metody filtrowania kolaboratywnego, szczególnie w początkowej fazie eksploatacji, jest niewielka ilość danych rozumianych jako ilość pustych komórek macierzy ocen/zdarzeń dla użytkowników. W celu eliminacji tego problemu podejmowane są próby wstępnej klasyfikacji użytkowników (przed zebraniem rzeczywistych danych mówiących o ich preferencjach) np. z wykorzystaniem tzw. zbiorów wskaźnikowych. W tym podejściu, każdy nowy użytkownik otrzymuje do oceny jeden stały zbiór elementów, na podstawie którego wyznaczana jest wstępna klasyfikacja użytkownika. Podejście to posiada wadę polegającą na tym, że zbiór wskaźnikowy musi być stały przez cały cykl życia systemu.

Zastosowanie profili w celu inteligentnego wyszukiwania

Proponowane rozwiązanie polega na grupowaniu użytkowników w klastry nie tylko ze względu na ich zachowanie ale na cechy profilu, który jest obowiązkowo wypełniany przy rejestracji. Z uwagi na charakter systemu wyszukiwania jest możliwe wpro-

wadzenie klasteryzacja na podstawie danych zawartych w profilu tj. w szczególności: wieku, płci, kraju (edukacji) i zainteresowań [30].

Podobnie jak w przypadku klasycznego filtrowania kolaboratywnego, tworzona jest macierz, ale w tym przypadku zawiera one cechy profili. Procedura postępowania jest zbliżona do procedury przy klasycznym filtrowaniu.

Propozycja systemu wyszukiwania spersonalizowanego

System spersonalizowany oparto o dwa niezależne podejścia do klasteryzacji. Pierwsza kategoria klastrów opisuje zachowanie użytkowników (w oparciu o collaborative filtering) podczas gdy druga uwzględnia informacje zdefiniowane przez użytkownika w swoim profilu. Każdy z klastrów nadrzędnych został dodatkowo podzielony na 3 podtypy opisujące zachowanie: nauczycieli, uczniów oraz łącznie wszystkich użytkowników zarejestrowanych.

Dodatkowo wprowadzono wyliczanie dopasowanie konkretnego obiektu do zadeklarowanych zainteresowań danego użytkownika.

Ostateczne dopasowanie stanowi wypadkową powyżej przedstawionych modeli, a parametry wpływające na wagę poszczególnych wartości mogą być modyfikowane przez użytkownika [11].

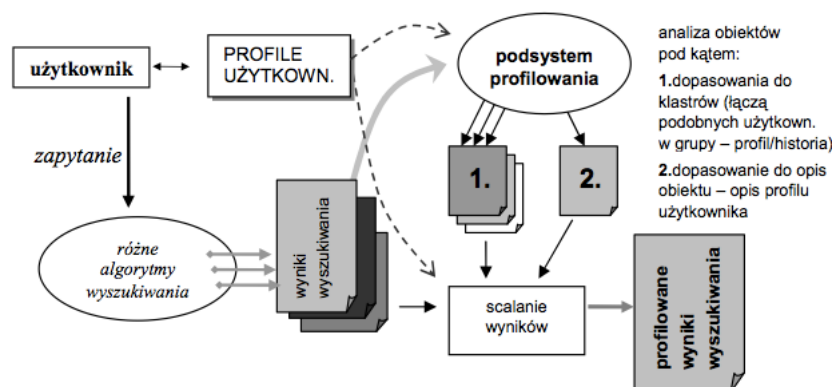
Zaprojektowany system analizy zachowań użytkownika wykorzystuje następujące podejścia:

- explicit obejmujących aktywne działanie użytkownika takie jak: rating (ocena) czy labelling (przypisanie etykiety, czy dodanie do ulubionych);
- implicit obejmuje analizę czynności podejmowanych przez użytkownika takich jak: selection (zaznaczenie wyników), viewing (obejrzenie szczegółowej informacji o zasobie) czy downloading (pobranie zasobów)

Ponadto w odróżnieniu do wielu systemów CF bazujących tylko na analizie czynności podejmowanych przez użytkowników zaprojektowany system wykorzystuje również analizę podobieństw pomiędzy użytkownikami na podstawie danych wpisanych w posiadany profil użytkownika. Analiza podobieństw wykorzystuje następujące elementy profilu: wiek, płeć, kraj edukacji and zdefiniowane zainteresowania. System analizy danych w profilu jest podobny do procedury używanej w „klasycznym” podejściu CF. Tworzona jest macierz $m \times n$, w której kolumny reprezentują wszystkie możliwe wartości każdego z 4ch analizowanych elementów profilu. Wprowadzono również metodę oceny wzajemnych relacji pomiędzy poszczególnymi składowymi poprzez uwzględnienie wag [30].

Propozycja metody tworzenia spersonalizowanej listy wyników

Proces wyszukiwania spersonalizowanego w zaprojektowanym systemie przebiega kilkietapowo (Rys. 5).



Rysunek 5. Graficzne przedstawienie procesu tworzenia personalizowanej listy wyników

1. W pierwszej fazie mają miejsce dwa zdarzenia: **wyszukanie** obiektów dydaktycznych spełniających zadane przez użytkownika kryteria w oparciu o zaimplementowane wyniki, a następnie ich **scalanie** w oparciu o pre-definiowane przez użytkownika ustawienia.

2. Faza druga polega na przekazaniu wartości dopasowań poszczególnych obiektów których ciąg uzyskano w fazie pierwszej w zależności od wybranego algorytmu profilującego.

W przypadku fazy drugiej celowe jest uzyskanie spójnej listy wartości uwzględniającej poszczególne wartości dopasowania. Przyjęto wykorzystanie średniej ważonej, a dzięki zebranych opiniom od użytkowników przyjęto wstępnie, że waga będzie liczona zgodnie ze wzorem:

$$W_i = \begin{cases} (a_i - b_i) \cdot c_i & \text{gd}y a_i \geq b_i \\ 0 & \text{gd}y a_i < b_i \end{cases} \quad (3)$$

gdzie: w oznacza wagę danego algorytmu, a oznacza procentową liczbę głosów powyżej oceny 3, b oznacza liczbę ocen poniżej 3, zaś c oznacza procentowe zainteresowanie danym algorytmem. Wartości a, b i c pochodzą z ankiet użytkowników oceniających prototyp systemu.

3. Faza trzecia jest opcjonalna i polega na (zgodnie z rysunkiem 5) połączeniu wartości dopasowań dla metod wyszukiwujących (uzyskanych jako wynik fazy 1) z wartościami dopasowań algorytmów profilujących (uzyskanych jako wynik fazy 2). Z uwagi na różny charakter wartości przyjęto zastosowanie średniej geometrycznej, a w uproszczeniu wyliczanie wartości zgodnie ze wzorem:

$$v_i = \frac{v_{q_i} \cdot v_{p_i}}{100} \quad (4)$$

gdzie v_i oznacza wynikową wartość dopasowania, v_{q_i} wartość dopasowania do metod wyszukiwawczych, a v_{p_i} wartość uzyskaną z podsystemu profilowania.

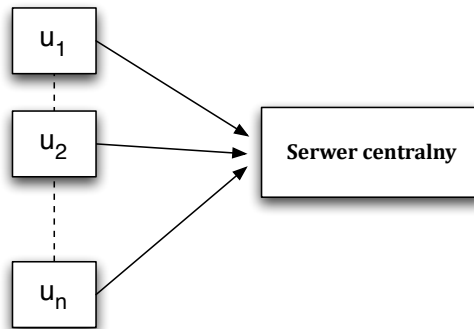
5. Analiza obciążenia systemów rozproszonych

Celem niniejszego podrozdziału jest przeprowadzenie analizy obciążenia systemu wielo-dostępowego, a zwłaszcza obciążenia serwera centralnego do którego ma dostęp wielu użytkowników.

Analizę przeprowadzono z zastosowaniem wybranych rezultatów strumieni impulsowych [10] oraz teorii masowej obsługi [16].

Przyjęcie określonych założeń upraszczających pozwala na otrzymanie konstruktywnych wyników, które według wiedzy autora, nie były przedstawiane dotychczas w dostępnej literaturze w odniesieniu do analizowanych w pracy rozproszonych systemów wyszukiwania informacji.

Przedmiotem analizy jest w ogólnym przypadku system przedstawiony na rysunku 6.



Rysunek 6. Schemat jednoczesnych połączeń z serwerem przez wielu użytkowników.

Wprowadzono w pracy pojęcie współczynnika **obciążenia serwera** przez i-tego użytkownika, który jest równy iloczynowi średniej liczby połączeń i-tego użytkownika z serwerem przez średni czas trwania takiego połączenia.

$$\eta_i = \bar{z}_i \bar{T}_i \quad (5)$$

gdzie:

\bar{z}_i średnia liczba połączeń z serwerem przez pojedynczego i-tego użytkownika.

\bar{T}_i średni czas trwania połączenia z i-tym użytkownikiem.

Wartość obciążenia serwera przez i-tego użytkownika zawiera się w przedziale:

$$\eta_i \in [0, 1]$$

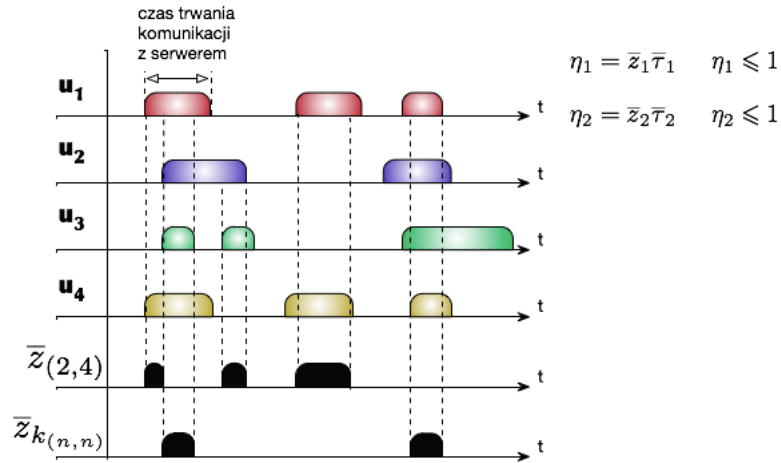
gdzie wartość 0 oznacza brak obciążenia zaś wartość 1 oznacza pełną zajętość serwera przez danego użytkownika.

Wprowadzono także następujące oznaczenia:

$\bar{z}_{k(n,n)}$ średnia liczba połączeń z serwerem przez wszystkich użytkowników.

$\bar{z}_{(j,n)}$ średnia liczba równoczesnych połączeń j-tej liczby użytkowników z całkowitej liczby n użytkowników korzystających z systemu.

Schemat komunikacji użytkownika z serwerem ilustruje rysunek 7.



Rysunek 7. Schemat połączeń z serwerem przez czterech użytkowników w przykładowym okresie czasu.

Na rysunku zaznaczono przedziały określające czas trwania komunikacji z serwerem przez danych użytkowników (kolory: dla u_1 : czerwony, u_2 : fioletowy, u_3 : zielony i dla u_4 : żółty), oraz przedziały czasu w którym serwer obsługiwał wszystkich użytkowników a także tylko dwóch spośród wszystkich czterech użytkowników korzystających z serwera w danym okresie czasu.

Po przyjęciu założeń upraszczających:

1. Połączenia użytkowników z serwerem są reprezentowane przez losowy i stacjonarny strumień impulsów,
2. Użytkownicy są niezależni,
3. Średnia liczba połączeń każdego użytkownika z serwerem jest jednakowa:

$$\bar{z}_1 = \bar{z}_2 = \bar{z}_3 = \dots = \bar{z}_n = \bar{z}$$

4. Średni czas połączenia każdego użytkownika

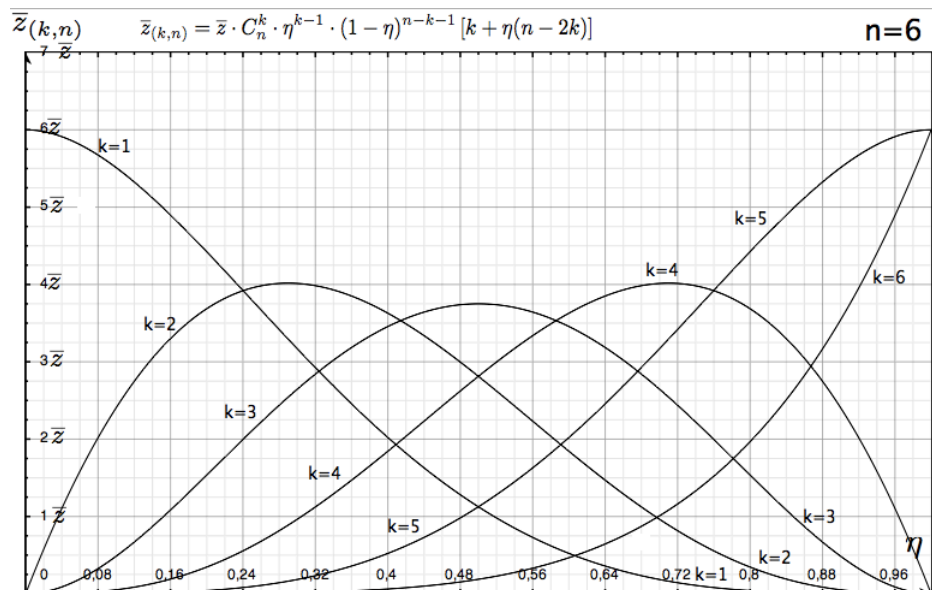
$$\bar{\tau}_1 = \bar{\tau}_2 = \bar{\tau}_3 = \dots = \bar{\tau}_n = \bar{\tau}$$

Uzyskano następujące wzory:

$$\bar{z}_{(k,n)} = \bar{z} \cdot C_n^k \cdot \eta^{k-1} \cdot (1 - \eta)^{n-k-1} [k + \eta(n - 2k)] \quad (6)$$

$$\bar{\tau}_{(k,n)} = (1 - \eta) \cdot \frac{1}{k + \eta(n - 2k)} \quad (7)$$

Na rysunku 8 przedstawiono wartość średniej liczby połączeń w funkcji wprowadzonego współczynnika obciążenia serwera na podstawie wzoru (6).

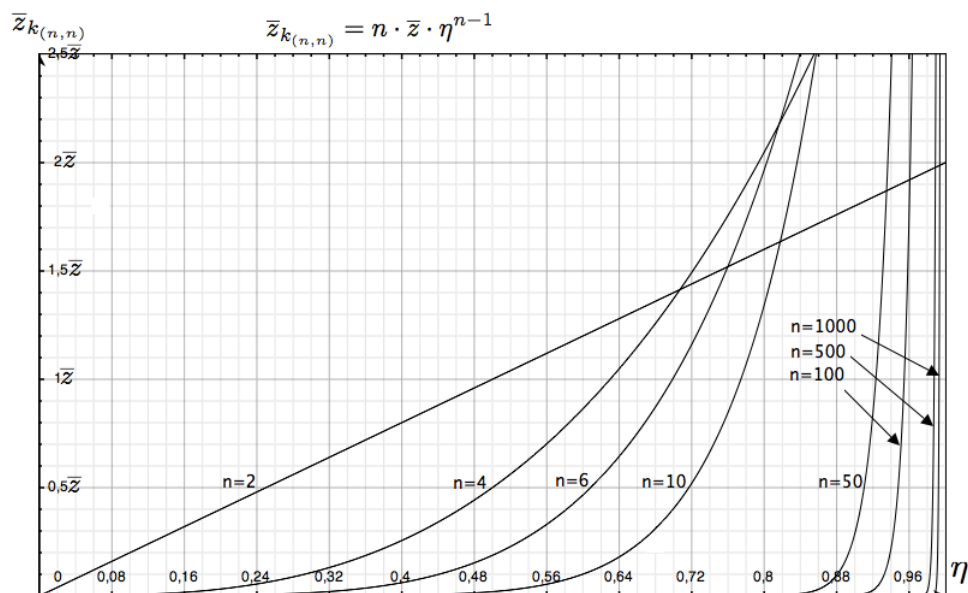


Rysunek 8. Wartości średniej liczby połączeń w funkcji współczynnika obciążenia serwera przy ustalonej całkowitej liczbie obsługiwanych użytkowników ($n=6$).

Na koniec przeanalizowano średnią liczbę połączeń z serwerem przez wszystkich użytkowników obsługiwanych przez serwer. Wzór (6) przyjmuje wtedy postać:

$$\bar{z}_{k(n,n)} = n \cdot \bar{z} \cdot \eta^{n-1} \quad (8)$$

Na rysunku 9 przedstawiono średnią liczbę połączeń przez wszystkich użytkowników przy różnej maksymalnej liczbie użytkowników obsługiwanych przez serwer.



Rysunek 9. Wartości średniego czasu zajęcia serwera przez wszystkich użytkowników.

Podane charakterystyki mogą w przyszłości posłużyć do jakościowej analizy wymiany informacji w systemach rozproszonych, zwłaszcza w aspekcie projektowania połączeń serwera i łączy telekomunikacyjnych. Podstawą analizy ilościowej są równania: (6-8).

6. Agentowy System Wyszukiwania

W pracach autora rozprawy wykazano, że systemy agentowe z powodzeniem mogą zostać wykorzystywane dla realizacji i organizacji rozproszonego systemu wyszukiwania, gdzie mogą być wykorzystywane zarówno w procesie samego przeszukiwania zasobów [19, 20, 25, 29] jak też rozpraszania informacji w rozproszonym systemie baz danych [21-24].

Autor zaproponował i przedstawił różne strategie wyszukiwania zasobów w zależności od lokalizacji w ramach grupy zajmującej się e-edukacją [24]. Obserwowana z roku na rok potrzeba „lepszej” organizacji zasobów przy wzroście ilości danych stawała się coraz bardziej widoczna szczególnie w połączeniu ze wzrostem możliwości sprzętowych, rozwojem telekomunikacji i stale zwiększającej się liczbie osób zainteresowanych korzystaniem z ICT.

W wyniku dalszych prac autor rozprawy zaproponował wykorzystanie systemu agentowego do realizacji wyszukiwania obiektów wiedzy, wdrożenie tego modelu, a następnie rozszerzenie go o wyszukiwanie spersonalizowane w ramach rozproszonego systemu e-learningowego.

Proponowane podejście ma charakter nowatorski gdyż dotychczas systemy agentowe nie były szeroko stosowane w e-edukacji, podobnie jak wyszukiwanie spersonalizowane.

6.1. Podstawowe zadania systemu

Podstawowym celem Agentowego Systemu Wyszukiwania (ang. *Agent Based Search System*) jest zapewnienie użytkownikom edukacyjnym szybkiego dostępu do żądanej informacji.

W fazie analizy przyjęto za istotne, iż odnalezienie zasobów dydaktycznych będzie skuteczne tylko wtedy gdy znaleziona lista zasobów będzie dopasowana do potrzeb konkretnego użytkownika. Tym samym szybkość wyszukiwania należy rozumieć nie tylko jako natychmiastową prezentację danych lecz raczej jako szybkość dotarcia przez użytkownika do zasobów, najistotniejszych z jego punktu widzenia. Przeprowadzone badania [31, 46] potwierdzają, że osoba korzystająca z systemów wyszukiwania w szczególności zwraca uwagę na początkowe elementy na liście wyników. Równocześnie czas spędzany na analizie znalezionych wyników nie jest wielki. Informacje te czynią istotnym pozycjonowanie znalezionych wyników.

Koncepcje wyszukiwania

Odnalezienie konkretnego zasobu wiedzy (LO) sprowadza się do znalezienia informacji o tym obiekcie i następnie zaprezentowanie / dostarczenie tej informacji do użytkownika. Każdy LO opisany jest zgodnie ze schematem konkretnego systemu (w tym przypadku jest LRE Application Profile) który to opis (LOM) wywodzi się ze standardu IEEE LOM. Aby zapewnić skuteczne wyszukiwanie danych opis (LOM) winien być dostępny dla zaproponowanego zorientowanego na agentów systemu wyszukiwania. Bazując na schemacie Federated Search zaproponowano możliwości wzbogacenia wyszukiwania w sieci niezależnych repozytoriów. Założono, że użytkownik uzyskuje dostęp i możliwość wyszukiwania czy przeszukiwania innych repozytoriów poprzez swój lokalny

system. W kolejnym kroku rozważań wprowadzono możliwość istnienia dwóch strategii:

- z serwerem zarządzającym wyszukiwaniem ;
- bez serwera zarządzającego wyszukiwaniem.

Dla każdej strategii przyjęto dodatkowe możliwości i tak: **wyszukiwanie z serwerem zarządzającym** może być realizowane poprzez:

(1) bezpośredni dostęp użytkownika do tego serwera, zlecenie wyszukania i oczekiwanie na rezultat, (2) ewentualnie serwer jest „przesłonięty” lokalnymi zasobami, a osoba nie ma do niego bezpośredniego dostępu. Realizacja wyszukania w tej strategii jest analogiczna z istniejącym podejściem w ARIADNE.

Z kolei dla **wyszukiwania bez serwera zarządzającego** rozważono 2 warianty: (3) użytkownik korzysta z aplikacji lokalnie, wysyła agenta przeszukującego zasoby i czeka na jego powrót (por. [29]), bądź (3') użytkownik zleca realizację wyszukania interesujących go zasobów poprzez lokalny system wyszukiwania zlokalizowany w jego intranecie. Model przedstawiony jako 3' w istocie jest podobny do modelu z serwerem zarządzającym którego część funkcjonalności jest realizowana po stronie jednostki centralnej w danym intranecie (por. "System designed for research centres" i "System designed for Internet users" w pracy autora rozprawy [19]).

Rozważono również dwie strategie realizacji zapytań pomiędzy repozytoriami realizowanymi przez agentów mobilnych (zaproponowane przez autora rozprawy [24]). Model pierwszy (określony jako "**each-others approach**") polega na wysyłaniu agentów z każdego repozytorium (3') bądź serwera centralnego (1,2) do każdego z repozytoriów w poszukiwaniu danych. Model drugi (określony jako "**circle approach**") (2,1') w którym agenci wędrują po sieci, a dane repozytorium może zlecić im wykonanie poszczególnych zadań (por. pojęcie „wolny agent” [6]).

Ostatecznie z uwagi na ograniczone możliwości dostępu do repozytoriów przyjęto model o możliwie najbardziej szerokiej funkcjonalności w celu łatwości jego adaptacji i dalszej rozbudowy, tj. wybrano model z jednostką centralną zajmującą się gromadzeniem metadanych pochodzących z repozytoriów w celu lokalnej ich analizy. Agenci mogą być wysyłani do repozytoriów wspierających środowisko agentowe gdzie będą odpowiadać za aktualizację kodu, oraz stale monitorować zmiany w danych hoście.

W przyjętym modelu system agentowy zarządza komunikacją z użytkownikiem końcowym poprzez osobistego asystenta (ang. *personal assistant*) i w imieniu użytkownika prowadzi wyszukiwanie i przetwarzanie informacji. System taki również można łatwo rozproszyć pomiędzy kilka maszyn w przypadku wzrostu jego obciążenia – liczby obsługiwanych użytkowników, czy przetwarzanych metadanych.

Dostęp do zasobów repozytoriów obiektów wiedzy LOR realizowany jest w oparciu o implementację SQI [38, 43, 44]. Przy czym przy czym format zapytania może być oparty o często stosowany S²QL (ang. *Simple School Query Language*) stanowiący rozszerzoną postać VSQL (ang. *Very Simple Query Language*). Rozszerzenie dotyczy możliwości wyszukiwania zasobów według wieku uczniów (age range) oraz języka w którym utworzono obiekty LO.

S²QL jest językiem zapytań opracowanym na potrzeby wyszukiwania treści dydaktycznych w repozytoriach. Nie jest związany ze strukturą danych i zapytania są lokalnie konwertowane do formatów zapytań obsługiwanych przez bazy danych (SQL).

Zapytania S²QL mają format XML zgodnie ze schemą XSD: fire.eun.org/xsd/s2ql-1.0.xsd

Pomimo faktu, iż SQI jest popularnym interfejsem dostępu do repozytoriów wydajniejsze w proponowanym systemie jest zastosowanie: OAI-PMH. Podstawowe ograniczenia przesądzające o rezygnacji z SQI dotyczą: ograniczeń w liczbie zwracanych wyników oraz braku możliwości przeglądania zawartości repozytoriów np. wg daty. Implementacje powszechnie stosowane wspierają tylko wyszukiwanie po słowach kluczowych.

Protokół OAI-PMH [41] jest realizowany w oparciu o XML i służy do selektywnego pobierania LOM. OAI-PMH umożliwia pobieranie danych zarówno w oparciu o słowa kluczowe jak i dat modyfikacji co jest przesądzające w przypadku systemu wyszukiwania obiektów dydaktycznych.

6.2. Analiza szczegółowa Systemu

System złożony jest z dwóch podstawowym modułów: modułu gromadzenia danych i modułu zaawansowanego wyszukiwania.

6.2.1. Moduł gromadzenia danych — Data Collection Module (DCM)

DCM jest modułem agentowego systemu wyszukiwania odpowiedzialnym za nadzorowanie stanu metadanych w repozytoriach, aktualizację danych do centralnego repozytorium metadanych ABSS oraz wstępną analizę zasobów [3, 12].

Rdzeniem DCM jest platforma agentowa powstała w oparciu o JADE-LEAP [5].

Przeanalizowano trzy scenariusze pobierania danych:

- w oparciu o dane uzyskane z systemu brokerskiego (Brokerage System) i komunikację przez sieć w oparciu o protokół OAI-PMH;
- poprzez platformę agentową i wprowadzoną tam obsługę agentów (LMS service) oraz komunikację lokalną po stronie repozytorium w oparciu o OAI-PMH;
- jak wyżej ale w oparciu o serwis SparkOAI.

6.2.2. Moduł Zaawansowanego Wyszukiwania (ASM)

Wyszukiwanie stosowane w rozproszonych systemach e-learningowych czy służących e-edukacji takich jak Ariadne, CELEBRATE, itp. oparte są o model Federated Search (SQI) który nie uwzględnia przy wyszukiwaniu profili użytkowników a tylko słowa kluczowe (VSQL) ew. dodatkowo wiek i język (S2QL).

W skład modułu zaawansowanego wyszukiwania ASM wchodzi dwa podsystemy które umożliwiają prawidłową realizację zaawansowanego wyszukiwania oferując profilowanie dostarczonych wyników wyszukiwania. Podsystemy te to: **podsystem personalizujący** przetwarzający profile użytkowników oraz **podsystem wyszukujący**. Elementem centralnym integrującym oba podsystemy jest agent osobisty stanowiący osobistego asystenta użytkownika.

Podsystem personalizujący gromadzi i przetwarza informacje o użytkownikach i czynnościach przez nich wykonywanych (np. listę pobranych, ocenionych obiektów) stale aktualizując klastry służące filtrowaniu wyników (w oparciu o klasyczny i zmodyfikowane podejście CF) oraz zapewnieniu rekomendacji (promowaniu) nowych wyników.

Podsystem wyszukujący dostarcza ukierunkowane na języki naturalne określone metody wyszukiwania w metadanych opisujących obiekty wiedzy LO oraz zapewnia

możliwość poszerzenia wyszukiwań poprzez zastosowanie m.in. słowników synonimów.

Podstawowymi elementami tworzącymi ASM są agenci przetwarzający dane zgromadzone w zarządzanych przez nich bazach danych.

6.3. Realizacja praktyczna systemu i opinie użytkowników

W trakcie prac przeprowadzono spotkania z nauczycielami celem zebrania opinii o użyteczności systemu we wspomaganiu wyszukiwania informacji. Zebrane opinie pozwoliły wdrożyć sugestie konkretnych osób w celu zapewnienia wysokiej skuteczności działania systemu wobec oczekiwań docelowej grupy użytkowników.

Przeprowadzone końcowe ankiety wskazują na stosunkowo duże zainteresowanie wyszukiwaniem spersonalizowanym. Z uwagi na fakt, iż zaimplementowany system spersonalizowany wymaga nauczenia się ocena przeprowadzona w krótkim czasie po implementacji nie oddaje w pełni jego jakości którą można uzyskać stosując system przez dłuższy czas. Jednak nawet pomimo tego faktu oceny wyszukiwania spersonalizowanego są dość wysokie, szczególnie w stosunku do klasycznego podejścia, co pozwala pozytywnie ocenić kierunek który został wyznaczony w pracy.

7. Podsumowanie

Rozwój systemów e-learningowych wymusza tworzenie coraz doskonalszych narzędzi pozwalających użytkownikom zarówno tworzyć jak i dzielić się obiektami wiedzy (LO).

W celu zapewnienia osobom tworzącym kursy dostępu do całego spektrum kursów czy elementarnych ich składowych tworzone są rozproszone sieci repozytoriów (LORs), zaś w celu zapewnienia łatwego dostępu do zasobów istnieją one w parze (logicznie) z repozytoriami zawierającymi opisy obiektów w formie metadanych (LOM). Struktury te przede wszystkim są zarządzane centralnie (np. ARIADNE – Federacja Repozytoriów, MACE – Super-repozytorium) co ma gwarantować łatwość i szybkość wyszukania (Federated Search) oraz następnie zapewnić szybki dostęp do znalezionych zasobów.

Powstają również inicjatywy mające na celu zapewnić łatwość i szybkość wymiany danych oraz integrację systemów LORs (np. GLOBE). Tym samym kluczową staje się potrzeba zapewnienia dostępu użytkownikom do konkretnych zasobów które zarówno są wysokiej jakości jak i najlepiej odpowiadają zainteresowaniom konkretnego użytkownika.

Należy stwierdzić, że cel pracy został zrealizowany i w rezultacie zaproponowano i przeanalizowano system zaawansowanego wyszukiwania bazujący na różnych aspektach personalizacji takich jak: podobieństwo pomiędzy użytkownikami (tradycyjne metody CF), podobieństwa w grupie użytkowników z profilami (modified-CF), podobieństwa pomiędzy obiektem, a konkretnym profilem użytkownika.

Analizowany system został zaprojektowany i zaimplementowany jako rozszerzalny moduł mogący być wykorzystywany wewnątrz platform eLearningowych.

Do najważniejszych rezultatów niniejszej pracy należy zaliczyć:

- Zaproponowanie i wdrożenie platformy agentowej do realizacji wyszukiwania zasobów w systemach e-learningowych opartych na rozproszonych repozytoriach (w tym tworzących sieć GLOBE), a w szczególności:

- zapewnienie bezpiecznej komunikacji pomiędzy rdzeniem systemu i zdalnymi repozytoriami;
- uniezależnienie systemu pobierania danych z repozytoriów od lokalnych operacji w tym restart głównego kontenera – rdzenia platformy agentowej;
- wprowadzenie wyspecjalizowanych agentów do realizacji określonych zadań od nadzorowania repozytoriów, obsługę metod wyszukiwania po wprowadzenie osobistego asystenta (personal agent) realizującego proces wyszukiwania w imieniu konkretnego użytkownika;
- wprowadzenie możliwości łatwej rozbudowy systemu poprzez dodawanie kolejnych wyspecjalizowanych agentów (np. przy wdrożeniu innych metod wyszukiwania czy pozycjonowania) oraz łatwego rozproszenia systemu z uwagi na wykorzystaną jego strukturę;
- Wprowadzenie obsługi wyszukiwania spersonalizowanego w tym:
 - zapis historii ich działań, oraz zaimplementowanie jawnego i pasywnego systemu oceniania znalezionych danych;
 - zaimplementowanie poza standardowymi metodami klasteryzacji opartymi o CF własnych modeli klasteryzacji użytkowników.
 - wprowadzenie modułu rankingowego pozwalającego na automatyczną i przezroczystą z punktu widzenia użytkownika personalizację wyników wyszukiwania.
 - zaproponowanie klasteryzacji na podstawie profili skracającej czas uczenia się przez system, preferencji nowych użytkowników.
- Wykonanie analizy obciążenia systemów rozproszonych dla przypadku ogólnego którą można łatwo zaadoptować dla konkretnych zadanych czynników

Przeprowadzone testy w pełni potwierdzają korzyści oferowane przez niniejszy System, a w szczególności:

- przetwarzanie zapytań może być realizowane w trybie off-line;
- system może być w łatwy sposób rozpraszany pomiędzy wiele maszyn w sieci przy wzroście ilości danych i liczby użytkowników;
- modyfikacja poszczególnych serwisów może być dokonywana w czasie rzeczywistym bez potrzeby wyłączania całego Systemu;
- wielopoziomowa analiza zachowań użytkowników w połączeniu z zadeklarowanymi informacjami pozwala w lepszy sposób dopasować rezultaty wyszukiwania do realnych oczekiwań konkretnych użytkowników.

W niniejszej pracy wykazano, że platformy wieloagentowe z powodzeniem nadają się do tworzenia systemu inteligentnego wyszukiwania obiektów dydaktycznych w ogólnodostępnych repozytoriach metadanych. Przedstawione cechy w pełni potwierdzają zalety stosowania technologii agentowej dla uzyskania wydajnego i inteligentnego modułu wyszukiwania spersonalizowanego w systemach e-Learningowych.

Wprowadzone korzyści oferowane przez wyszukiwanie spersonalizowane zostało docenione przez użytkowników końcowych, a korzyści oferowane będą widoczne jeszcze mocniej kiedy system zostanie w pełni skalibrowany co wymaga jednak dłuższego czasu. Wstępna kalibracja oparta na wynikach ankiet okazała się jednak już na etapie walidacji celowa, a korzyści zostały dostrzeżone nie tylko przez użytkowników końcowych lecz i ekspertów dziedzinowych.

Zastosowanie powyższego modułu pozwala na uzyskanie inteligentnego wyszukiwania informacji. Znalezione obiekty LO które spełniają kryteria użytkownika są następnie przetwarzane i odpowiednio pozycjonowane. Kolejność wyników jest wypadkową analizy zainteresowań użytkownika (zadeklarowanych i realnych), popularności i jakości danego LO oraz zgodności obiektu z zadeklarowanymi zainteresowaniami.

Wyniki prac autora rozprawy zostały opublikowane w formie rozdziału p.t. "The Use of Multi-Agents' Systems in e-Learning Platforms" w książce "E-learning", ISBN 978-953-7619-23-7, Wyd. IN-TECH w 2010 roku.

Bibliografia

- [1] G. Adomavicius, and A. Tuzhilin, **Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions**, *IEEE transactions on knowledge and data engineering*, 17, 2005, pp. 734-749.
- [2] L. Anido, **An observatory for e-learning technology standards**, *Advanced Technology for Learning*, 3, 2006, ACTA Press, pp. 99-108.
- [3] R. Baran, et al., **The data collection module of the agent based search system**, *Proc. of WEBIST'2007: "WEB INFORMATION SYSTEMS AND TECHNOLOGIES"*, March 3-6 2007, Barcelona, Spain,
- [4] F. Bellifemmine, A. Poggi, and G. Rimassa, **Jade – A FIPA compliant Agent Framework**, 1999,
- [5] G. Caire, and F. Pieri, **LEAP user guide**, 2006, <http://jade.tilab.com/doc/LEAPUserGuide.pdf>
- [6] K. Cetnarowicz, **Problemy projektowania i realizacji systemów wieloagentowych**, 80, AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków, 1999.
- [7] Chmiel, M. Gawinecki, P. Kaczmarek, M. Szymczak, and M. Paprzycki, **Efficiency of JADE agent platform**, *Proc. of International Symposium of Parallel and Distributed Computing & International Workshop on Algorithms, Models and Tools for Parallel Computing on Heterogenous Networks*, vol.13(2), April 2005, Amsterdam, The Netherlands, IOS Press, pp. 159-172.
- [8] G. Dobrowolski, **Technologie agentowe w zdecentralizowanych systemach informacyjno-decyzyjnych** (Polish), *Rozprawy Monograficzne 107*, Kraków, AGH – Uczelniane Wydawnictwa Naukowo-Dydaktyczne, 2002 ISSN: 0867-6631
- [9] E. Duval, **Standardized Metadata for Education: a Status Report**, *Proc. of the AACE World Conference on Educational Multimedia, Hypermedia and Telecommunications, ED-MEDIA'2001*, 2001, Norfolk, VA, AACE, pp. 458-463.
- [10] A. Dziech, **Random Pulse Streams and their Applications**, Elsevier, 1993.
- [11] A. Dziech, and T. Orzechowski, **Zastosowanie systemów agentowych w platformach e-learningowych**, *Przegląd Telekomunikacyjny*, 8-9, 2007, Stowarzyszenie Elektryków Polskich, pp. 808-813
- [12] A. Dziech, and T. M. Orzechowski, **The Agent-Base Search System**, *Proc. of the 2008 International Joint Conference on e-Commerce, e-Administration, e-Society, and e-Education. e-CASE'2008*, March 27-29 2008, Bangkok, Thailand
- [13] D. Goldberg et al., **Using collaborative filtering to weave an information tapestry**, *Communications of the ACM*, 35, 1992, ACM, pp. 70.
- [14] Informacje dot. wykorzystania e-learningu przez Firmę IBM, <http://www.304.ibm.com/jct03001c/services/learning/ites.wss/pl/pl?pageType=page&c=a0004008>
- [15] JADE, Java Agent DEvelopment Framework, Strona projektu zawiera dokumentację oraz w pełni funkcjonalne kody binarne i źródłowe platformy wieloagentowej <http://jade.tilab.com>
- [16] L. Kleinrock, and R. Gail, **Queueing Systems: Problems and Solutions**, John Wiley & Sons, 1996.
- [17] M. Montaner, B. López, and J. L. de la Rosa, **A Taxonomy of Recommender Agents on the Internet**, *Artificial Intelligence Review*, 19, 2003, Kluwer Academic Publishers, pp. 285-330.

- [18] C. O' Riordan, and H. Sorensen, **Information Filtering and Retrieval: An Overview**, *Technical Report*, 2003, Department of Information Technology, NUI Galway, Ireland,
- [19] T. M. Orzechowski, **Concepts of systems for information retrieval**, *Proc. of the IEEE Siberian Workshop of Students and Young Researchers, Modern Communication Technologies, SIBCOM'2001.*, November 28–29 2001, Tomsk, Russia, pp. 39-44.
- [20] T. M. Orzechowski, **Systemy agentowe w rozproszonych systemach zarządzania**, *AUTOMATYKA*, 5, 2001, AGH - Uczelniane Wydawnictwa Naukowo-Dydaktyczne, pp. 465-477.
- [21] T. M. Orzechowski, **The system for data exchange among databases oriented on CORBA standard and mobile agent technology**, *Proc. of the IEEE Siberian Workshop of Students and Young Researchers, Modern Communication Technologies, SIBCOM'2001.*, November 28–29 2001, Tomsk, Russia, pp. 14-19.
- [22] T. M. Orzechowski, **Voyager as a powerful tool for distributed programming**, *Proc. of the IEEE Siberian Workshop of Students and Young Researchers, Modern Communication Technologies, SIBCOM'2001.*, November 28–29 2001, Tomsk, Russia, pp. 30-34.
- [23] T. M. Orzechowski, **Meta-data distribution in intelligent databases system**, *Proc. of the 3rd IEEE Annual Siberian Workshop on Electron Devices and Materials, EDM'2002*, vol.2, July 1-5 2002, Erlagol, Russia, pp. 21-24.
- [24] T. M. Orzechowski, **Updating system for distributed databases**, Cancun, Mexico, 2002, ACTA Press, pp. 421-425.
- [25] T. M. Orzechowski, **Ontology oriented systems of the access to information**, *Proc. of the IEEE-Siberian Conference on Control and Communications, SIBCON'2003.*, October 1-2 2003, Tomsk, Russia, pp. 3-9.
- [26] T. M. Orzechowski, **Simulation approach to MAO Service implementation.**, *Proc. of the IEEE International Conference "Computer as a Tool", EUROCON'2005.*, vol.1, November 22-24 2005, Belgrade, Serbia and Montenegro, pp. 684-687.
- [27] T. M. Orzechowski, **The usage of mobile agents for management of data transmission in telecommunications' networks**, *The IEEE International Siberian Conference on Control and Communications, SIBCON'2005.*, 2005, pp. 8-12.
- [28] T. M. Orzechowski, **Zastosowanie agentów mobilnych do realizacji usług w sieciach telekomunikacyjnych**, Kraków, Poland, 2005, pp. 499-501.
- [29] T. M. Orzechowski, and A. Dziech, "COMMUNICATION BETWEEN END-USERS AND DATABASES USING MOBILE AGENT TECHNOLOGY," Honolulu, Hawaii, USA, 2001, ACTA Press, pp. 433-436.
- [30] T. M. Orzechowski, S. Ernst, and A. Dziech, "Profiled Search Methods for e-Learning Systems," Crete, Greece, 2007,
- [31] B. Pan, et al., "In Google we trust: Users' decisions on rank, position, and relevance," *Journal of Computer-Mediated Communication*, 12, 2007, <http://jcmc.indiana.edu/vol12/issue3/pan.html>
- [32] M. J. Pazzani, **A framework for collaborative, content-based and demographic filtering**, *Artificial Intelligence Review*, 13, 1999, Springer, pp. 393-408.
- [33] Portal poświęcony e-learningowi: <http://elearning.pl>
- [34] G. Salton, and C. Buckley, **Term-Weighting Approaches in Automatic Text Retrieval.**, *Information Processing & Management*, 24(5), Elsevier 1988
- [35] G. Salton, A. Wong, and C. S. Yang, **A Vector Space Model for Automatic Indexing**, *Communications of the ACM*, 18(11) 1975, pp. 613-620.
- [36] G. Salton, and M. J. McGill, **Introduction to Modern Information Retrieval**, New York, USA, McGraw-Hill Inc, 1986
- [37] E. Shakshuki, and Y. Jun, **Multi-Agent Development Toolkits: An Evaluation**, *Lecture Notes in Computer Science (Innovations in Applied Artificial Intelligence)*, 3029, 2004, Springer Berlin, pp. 209-218.
- [38] B. Simon, D. Massart, F. van Assche, S. Ternier, E. Duval, S. Brantner, D. Olmedilla, and Z. Miklós, **A Simple Query Interface for Interoperable Learning Repositories**, *Proc. of the 1st Workshop on Interoperability of Web-based Educational Systems*, May 2005, Chiba, Japan, pp. 11-18.

- [39] S. Sokołowski, and E. Sokołowska-Katzer, **Idee McLuhana inspiracją w modernizacji nauczania na odległość** (Polish), http://elearning.pl/filespace/artykuly/Sokolowski_3.pdf
- [40] Standard: **IEEE 1484 Learning Objects Metadata (IEEE LOM)**, 2005, http://tsc.ieee.org/wg12/files/IEEE_1484_12_03_d8_submitted.pdf
- [41] Standard: **The Open Archives Initiative Protocol for Metadata Harvesting**, <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>
- [42] Strona Fundacji ARIADNE, <http://www.ariadne-eu.org>
- [43] S. Ternier, and E. Duval, **Interoperability of Repositories: The Simple Query Interface in ARIADNE**, *International Journal on E-Learning*, 5, 2006, AACE, pp. 161-166. <http://go.editlib.org/p/21734>
- [44] F. van Assche, et al., **Spinning Interoperable Applications for Teaching & Learning using the Simple Query Interface**, *Journal of Educational Technology & Society*, 9, 2006, International Forum of Educational Technology & Society, pp. 51-67.
- [45] F. van Assche, and D. Massart, **Federation and Brokerage of Learning Objects and Their Metadata**, Washington, DC, USA, 2004, IEEE Computer Society, pp. 316-320.
- [46] M. S. van Gisbergen, J. van der Most, and P. Aelen, **“Visual attention to Online Search Engine Results,”** http://www.checkit.nl/pdf/eyetracking_research.pdf
- [47] C. J. van Rijsbergen, **Information Retrieval**, Butterworths, London, 1979.
- [48] G. Vigna, **Mobile Agents: Ten Reasons For Failure**, Berkeley, California, USA, 2004, pp. 298-299.