

KONCEPCJA ELASTYCZNEGO SYSTEMU SIECIOWYCH USŁUG REKOMENDACYJNYCH

Andrzej Marciniak

Katedra Podstaw Techniki, Uniwersytet Przyrodniczy w Lublinie

Streszczenie. W pracy przedstawiono koncepcję elastycznego systemu sieciowych usług rekomendacyjnych. Usługi rekomendacyjne są rodzajem usług informacyjnych świadczonych również przez systemy doradztwa rolniczego. Zakłada się, że algorytm tworzący ranking rekomendowanych produktów będzie powstawał w wyniku maszynowego uczenia na zweryfikowanych danych. Aplikacja modelująca taki algorytm może być wykonana w technologii typowej dla sieciowych baz wiedzy. Problemem jest to, że aktualne systemy sieciowych usług informacyjnych są niewystarczająco odporne na zakłócenia w dostępie do danych, z których syntetyzowana jest usługa. Stąd pomysł uelastycznienia systemu poprzez semantyczną integrację zasobów sieciowych.

Słowa kluczowe: usługi rekomendacyjne, uczenie maszynowe, elastyczne systemy informacyjne, semantyczna integracja danych

Wstęp

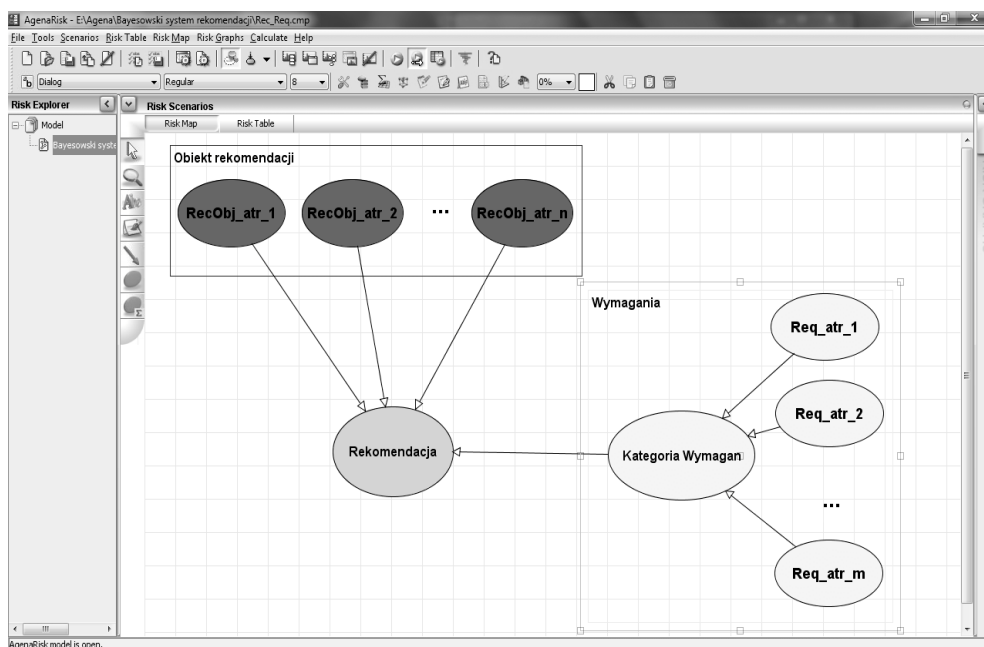
Celem pracy jest zbadanie możliwości budowania elastycznych, sieciowych systemów rekomendacyjnych. Pod pojęciem systemu rekomendacyjnego rozumiemy tu aplikację, która tworzy ranking obiektów z pewnego dobrze zdefiniowanego zbioru. Obiektami tymi mogą być, np. maszyny, nasiona, środki ochrony, nawozy ale także różnego rodzaju działania. Rozpatrywane obiekty reprezentowane są poprzez swoje atrybuty. Ranking obiektów oznacza utworzenie w ich zbiorze pewnego porządku. Ogólnie przyjmujemy że porządek taki definiujemy za pomocą relacji: *dominuje_{nad}*(O_i, O_j), co oznacza, że obiekt O_i poprzedza w rankingu obiekt O_j przy uporządkowaniu od najbardziej odpowiedniego do najmniej odpowiedniego w danej sytuacji. Sytuację tę definiuje kontekst w jakim przeprowadza się wartościowanie obiektów. Ogólnie założymy, że kontekst ten definiowany jest iteracyjnie przez klienta usługi rekomendacyjnej.

Istotnym założeniem pracy jest to, że algorytm odwzorowujący atrybuty wartościowanych obiektów i wymagania użytkownika w listę rankingową będzie powstawał w wyniku uczenia maszynowego na zweryfikowanych danych. Celem uczenia maszynowego jest wydobywanie tkwiących w danych ale nieznanych zależności i zapisanie ich w formie wykonywalnych modeli komputerowych, których podstawowym rodzajem wykonania jest inferencja wartości pewnego zbioru zmiennych w oparciu o zadane wartości pozostałych zmiennych. Model taki zachowuje się jak baza wiedzy udzielająca odpowiedzi na zadawane pytania.

Proponowane technologie

Wśród algorytmów spełniających wyżej sformułowane postulaty zaproponowano: sztuczne sieci neuronowe, systemy reprezentacji wiedzy oparte na logice, w tym systemy reguł heurystycznych, systemy optymalizacyjne, w tym systemy szukania oparte na algorytmach genetycznych oraz sieci probabilistyczne, bayesowskie lub oparte na teorii warunkowych pól losowych.

Na rysunku 1 przedstawiono sieć bayesowską generującą ranking wartościowanych obiektów. Wejście sieci tworzą dwie grupy węzłów – węzły reprezentujące atrybuty wartościowanych obiektów oraz węzły reprezentujące wymagania klienta usługi. Wyjście sieci stanowi grupa węzłów reprezentujących poszczególne obiekty.

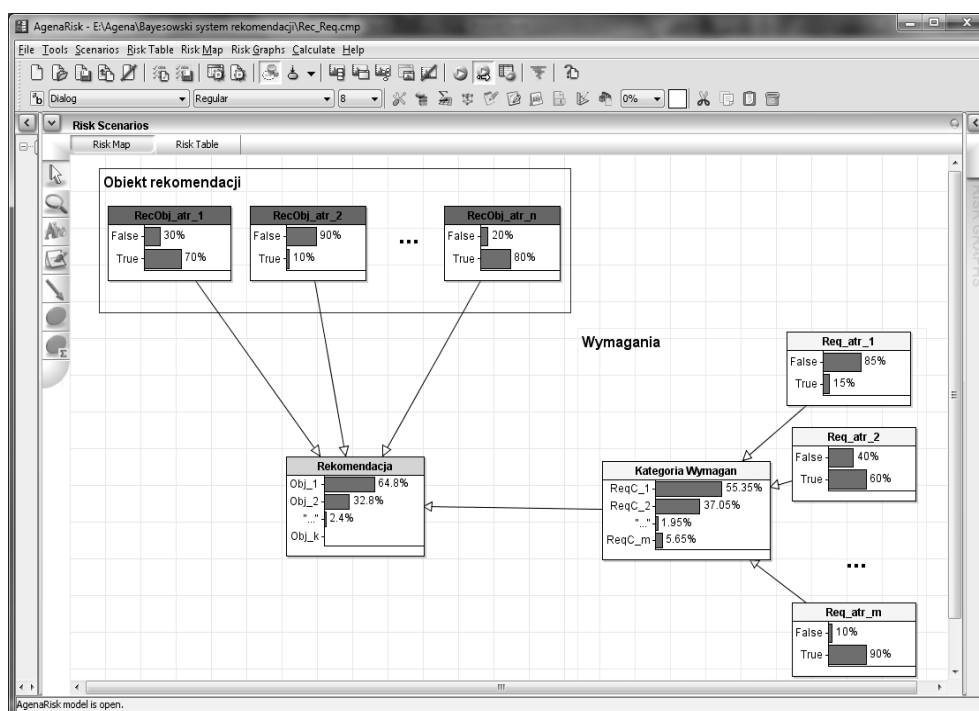


Rys. 1. Model systemu rekomendacyjnego w postaci sieci bayesowskiej

Fig. 1. Model of recommendation system in technology of Bayesian network

W pokazanym tu modelu zastosowano standardową dla technologii sieci bayesowskich binaryzację atrybutów wartościowanych obiektów. Przeprowadza się ją w ten sposób, że najpierw dyskretyzuje się zbiory wartości poszczególnych atrybutów (o ile już ze swej natury nie są one dyskretne), a następnie każdy atrybut o n wartościach zamieniamy na n atrybutów binarnych o wartościach $\{true, false\}$. Podobnie przeprowadzamy binaryzację wymagań klienta. Odpowiada to naturalnemu scenariuszowi pobierania informacji od klienta poprzez listę (sieć) pytań rozstrzygających, z których część jest od siebie niezależna a część jest zależna od odpowiedzi na wcześniej zadane pytania. Oznacza to, że wymagania użytkownika zostają sprecyzowane w wyniku sterowanego regułami przejścia przez sieć pytań.

W celu ograniczenia złożoności obliczeniowej była w tym modelu fuzja informacji reprezentowana przez zmienną *Kategoria_wymagań*. Na rysunku 2 pokazano rozkłady prawdopodobieństwa nad poszczególnymi zmiennymi modelu z rysunku 1. Rozkłady te odpowiadają sytuacji gdy atrybuty obiektów i wymagania klienta są „miękkimi faktami” (soft evidence). W modelach probabilistycznych, ranking ma postać rozkładu prawdopodobieństwa nad zbiorem obiektów. Pojęcie dominacji obiektu *i* nad obiektem *j* rozszerza się na przypadek dominacji probabilistycznej.

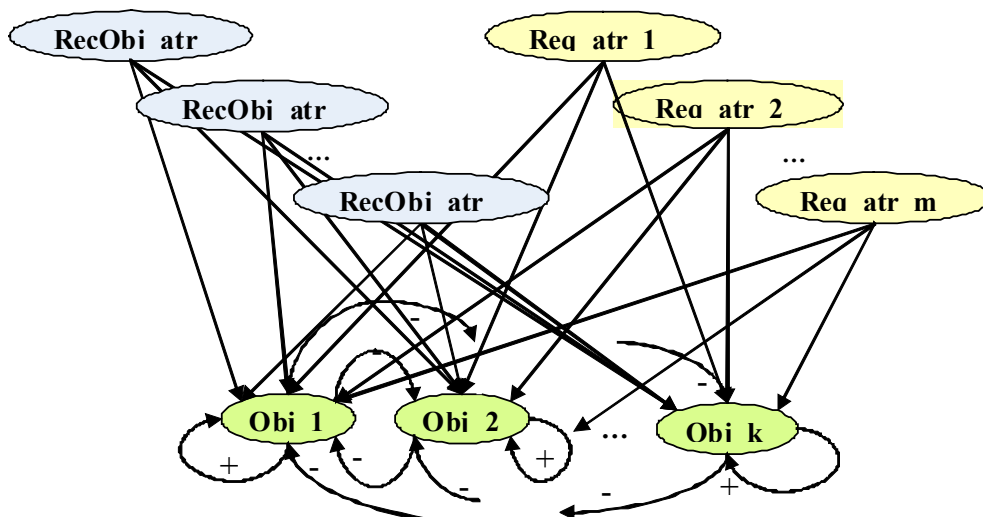


Rys. 2. Wynik rankingu jako rozkład prawdopodobieństwa nad wartościowanymi obiektami
Fig. 2. Result of the ranking as a distribution of probability for evaluated systems

Na rysunku 3 pokazano schemat systemu rekomendacyjnego opartego na sztucznej sieci neuronowej. Spośród wielu odpowiednich do tego architektur wybrano sieć Kohonena.

Wejściem tej sieci są atrybuty obiektów oraz wymagania klienta a wyjście tworzą kategorie reprezentujące poszczególne obiekty. W tym przypadku nie musimy podawać, w procesie uczenia sieci na przykładach, informacji o miejscu obiektów w rankingu.

Użyteczą konceptualizacją do budowy systemu rekomendacyjnego może być również ogólne sformułowanie problemu optymalizacyjnego w kategoriach algorytmów genetycznych. W tym przypadku zakładamy, że optymalny ranking obiektów to taki, który uwzględnia wymagania klienta i spełnia je w sposób maksymalnie możliwy. W metodzie tej, kody generowanej populacji rozwiązań problemu optymalizacyjnego (poszukiwane rankingi) ewoluują w wyniku selekcji, krzyżowania i mutacji.



Rys. 3. Sztuczna sieć neuronowa jako system tworzący ranking rekomendacyjny
 Fig. 3. Artificial neural network as a system creating the recommendation ranking

Elastyczne systemy sieciowe

W klasycznych aplikacjach sieciowych, dane są sztywno połączone z aplikacjami przetwarzającymi te dane. Systemy takie określamy jako „kruche” ponieważ zmiana adresu źródła danych lub najczęściej zmiana formatu danych powoduje, że aplikacja taka przestaje działać. Pojęcie elastyczności systemu odnosi się tu do jego niewrażliwości na takie zdarzenia. Konceptualizacja systemu elastycznego może tu być oparta na następujących stwierdzeniach:

- aplikacje nie są połączone sztywno z danymi,
- w poszukiwaniu danych, w sposób inteligentny przeszukiwane są globalne zasoby sieciowe,
- ażeby to było możliwe zasoby danych muszą być semantycznie adnotowane za pomocą odpowiednich metadanych,
- dane i metadane tworzą jedną spójną całość” [Segaran, T., *et al.* 2009].

Najprostszym sposobem semantycznej integracji danych są, np. bazy danych typu TripleStore. Ich architektura dokładnie odpowiada strukturze prostych asercji używanych w języku naturalnym – $\langle \text{subject} \rangle \langle \text{predicate} \rangle \langle \text{object} \rangle$. Część zdania $\langle \text{subject} \rangle$ (podmiot) reprezentuje byty (rzeczy), które w konceptualizacji tworzą zbiór podstawowych pojęć, $\langle \text{predicate} \rangle$ reprezentuje w najprostszym ujęciu własności bytów $\langle \text{subject} \rangle$, natomiast $\langle \text{object} \rangle$ to są bądź byty uczestniczące w innych relacjach, w których występują jako $\langle \text{subject} \rangle$ lub są literalami takimi jak *string* lub *number*. Oznacza to, że poprzez atrybut $\langle \text{object} \rangle$ relacji $\langle \text{predicate} \rangle$ mogą być ze sobą wiązane bardzo duże zbiory zdań (asercji) zachowując jednocześnie swoją prostą trójelementową strukturę. Modelem matematycznym takich

systemów są grafy skierowane (grafy pojęciowe), których węzły reprezentują byty pojęciowe

a łączące je łuki odpowiadają relacjom (predykatom). Dla takich repozytoriów danych można stosować algorytmy wyszukiwania w rodzaju „pattern matching” w których to o co pytamy reprezentowane jest jako zmienna umiejscowiona w pozycji określonego argumentu predykatu. Interesujące jest to, że tablicowe mechanizmy wnioskowania umożliwiają umieszczanie zmiennych również w miejscu nazwy predykatu. Reguły wnioskowania uruchamiane poprzez zapytania generują nowe asercje w formacie *(subject) (predicate) (object)*, a to oznacza, że jedne reguły mogą pracować na wynikach innych bez potrzeby jawnej ich koordynacji.

Oprócz dedukcyjnych mechanizmów rozumowania (produkowanie nowych asercji w oparciu o już istniejące) możliwe jest rozumowanie hipotetyczno-dedukcyjne sprowadzające się w tym przypadku do poszukiwania reguł nadających wartość zmiennym będącym bezpośrednim bądź pośrednim celem wnioskowania.

Nie ma tu istotnych ograniczeń w konstruowaniu złożonych łańcuchów inferencyjnych. Semantyczna integracja różnych zasobów danych sprowadza się do łączenia reprezentujących je grafów pojęciowych poprzez wspólne węzły, o ile reprezentowane przez nie pojęcia są jednoznacznie określone i znormalizowane. Semantyczna baza danych wyposażona w takie mechanizmy staje się silnie modułarna, odporna na błędy, łatwo dystrybuowalna i elastyczna.

Standardy semantyczne i formaty danych umożliwiające konstruowanie takich systemów już istnieją, [Obrst, Smith 2003; Grigoris, van Harmelen 2004; Watson 2010]. Np. system URI umożliwia tworzenie mocnych identyfikatorów a meta język RDF (<http://w3c.org/rdf>) opis bytów do których się one odnoszą. Język SPARQL jest standardem zapytań semantycznych dotyczących grafów RDF. Do danych i tworzenia zapytań można wykorzystać środowisko Java Framework o nazwie Sesame (<http://openrdf.org>).

Pewnym problemem jest fakt, że większość zasobów sieciowych nie jest utworzona z zachowaniem standardów semantycznych. Wyszukiwanie zasobów spełniających te standardy jest łatwe przy użyciu darmowego systemu Freebase i języka zapytań MQL (<http://www.freebase.com>). Istotnym krokiem w kierunku semantycznej reprezentacji zasobów sieciowych jest tworzenie w sieci repozytoriów ontologicznej reprezentacji wiedzy z zakresu konkretnych dziedzin przedmiotowych. Odpowiednim do tego standardem jest język OWL (<http://w3c.org/owl>) specjalnie przystosowany do definiowania klas i własności co umożliwia rozumowanie prowadzące do odpowiedzi na pytania co istnieje w danej dziedzinie, co jest czym i jakie ma własności. Formalna semantyka OWL określa logiczne konsekwencje zadanego w ontologii zbioru faktów, tj. nowe fakty literalnie w danej ontologii nieobecne. Te nowe fakty mogą być logiczną konsekwencją dynamicznie integrowanych, rozproszonych ontologii

Grafy są dobrym modelem samoopisujących się i spójnych chociaż rozproszonych zasobów sieciowych. Taką formę samorzutnie przybiera dzisiejszy globalny Web. Funkcjonuje nawet określenie „Global Graph Community” [Semantic Wave 2008].

Bibliografia

- Obrst L. J., Smith K. T.** 2003. The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management. J, Wiley. ISBN 0-471-43257-1
- Grigoris A., van Harmelen F.** 2004. A Semantic Web Primer. The MIT Press. ISBN-10: 0-262-01210-3, ISBN-13: 978-0-262-01210-2
- Segaran T., Evans C., Taylor J.** 2009. Programming the Semantic Web. O'Riley Media, Inc. ISBN: 978-0-596-15381-6
- Watson M.** 2010. Practical Semantic Web and Linked Data Applications, http://www.markwatson.com/opencontent/book_java.pdf
- Semantic Wave Report 2008. Project10X Semantic Wave research series. Dostępny w internecie: <http://project10x.com/>.

CONCEPTION OF THE FLEXIBLE SYSTEM OF NETWORK RECOMMENDATION SERVICES MODEL

Abstract. The work presents the conception of the flexible system of network recommendation services. Recommendation services are a kind of information services provided also by agricultural consultancy services. It is assumed that the algorithm creating the ranking of recommended products will be formed as a result of machine learning on the basis of verified data. The application modelling such an algorithm can be created in a technology which is characteristic of network knowledge bases. The problem is that current systems of network information services are insufficiently resistant to interruptions of access to data from which the service is synthesized. Consequently, an idea of making the system more flexible through semantic integration of network resources was created.

Key words: recommendation services, machine learning, flexible information systems, semantic data integration

Adres do korespondencji:

Andrzej Marciniak; e-mail: Andrzej.Marciniak@up.lublin.pl
Katedra Podstaw Techniki
Uniwersytet Przyrodniczy w Lublinie
Doswiadczalna 50A
20-280 Lublin