



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Wybrane problemy
Tom 11

Pod redakcją
Jerzego HOŁUBCA

Warszawa 2009



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Wybrane problemy
Tom 11

Pod redakcją
Jerzego HOŁUBCA

Warszawa 2009

Wykaz opiniodawców artykułów zamieszczonych
w niniejszym tomie:

prof. dr hab. inż. Jerzy HOŁUBIEC
dr inż. Lech KRUŚ
doc. dr hab. inż. Wiesław KRAJEWSKI
doc. dr hab. Jacek MALINOWSKI
dr inż. Edward MICHALEWSKI
prof. dr Adam SKOREK
dr hab. Ryszard SMARZEWSKI
prof. dr hab. inż. Andrzej STRASZAK
dr Dominik ŚLĘZAK
prof. dr hab. inż. Stanisław WALUKIEWICZ
doc. dr hab. Sławomir ZADROŻNY

© Instytut Badań Systemowych PAN
Warszawa 2009

ISBN 9788389475220

Druk: Zakład Poligraficzny Jerzy Kosiński, Warszawa

WYKRYWANIE WYŁUDZEŃ W UBEZPIECZENIACH KOMUNIKACYJNYCH

Piotr Dżidowski

Studia Doktoranckie IBS PAN

W artykule zaproponowano zastosowanie teorii zbiorów rozmytych do wykrywania wyłudzeń w ubezpieczeniach komunikacyjnych. Podany przykład ilustruje zalety proponowanego zastosowania.

Wstęp

Wyłudzenia ubezpieczeń to najczęściej popełniane przestępstwa ubezpieczeniowe. Jak podaje Polska Izba Ubezpieczeń (PIU), szacuje się że około 20-30 procent zgłoszeń zawiera elementy wyłudzeń, zaś średnią wartość wyłudzonych odszkodowań to około 20 tys. złotych. Dlatego też niezwykle istotnymi są w firmach ubezpieczeniowych komórki zajmujące się ich wykrywaniem. Zwykle są to zespoły ekspertów, które „ręcznie” analizują wybrane zgłoszenia. Metoda ta jest niestety niezwykle czasochłonna i obejmuje jedynie wybrane zgłoszenia, co może prowadzić do opóźnień w wypłacie należnego świadczenia oraz pominięcie szkód w których wystąpił element wyłudzenia. Dodatkowo większości firm nie zależy tylko i wyłącznie na wykryciu próby wyłudzenia, ale także na bliskim pewnością stwierdzeniu, że takiej próby w danym zgłoszeniu nie ma. Taka informacja pozwala na szybsze procesowanie takiego zgłoszenia.

Z tych powodów towarzystwa ubezpieczeniowe poszukują alternatywnych rozwiązań. Na rynku jest wiele firm świadczących usługi analizy danych i ich oceny pod kątem wykrywania wyłudzeń. Niestety w większości jakość ich rozwiązań nie jest zadowalająca, gdyż są one oparte głównie na analizie statystycznej. Dodatkowo usługi te są dość kosztowne i wahają się obecnie od około 2 do 4 euro za przeanalizowany rekord. W przypadku kilkudziesięciu tysięcy zgłoszeń miesięcznie daje to znaczną kwotę, która nie zawsze jest odzyskiwana w postaci niewypłaconia nienależnych odszkodowań.

Poniżej przedstawiono propozycję systemu który nie opierałby się na analizie statystycznej, a na analizie danych zbieranych na każdym etapie likwidacji szkody wykorzystującej zbiory rozmyte.

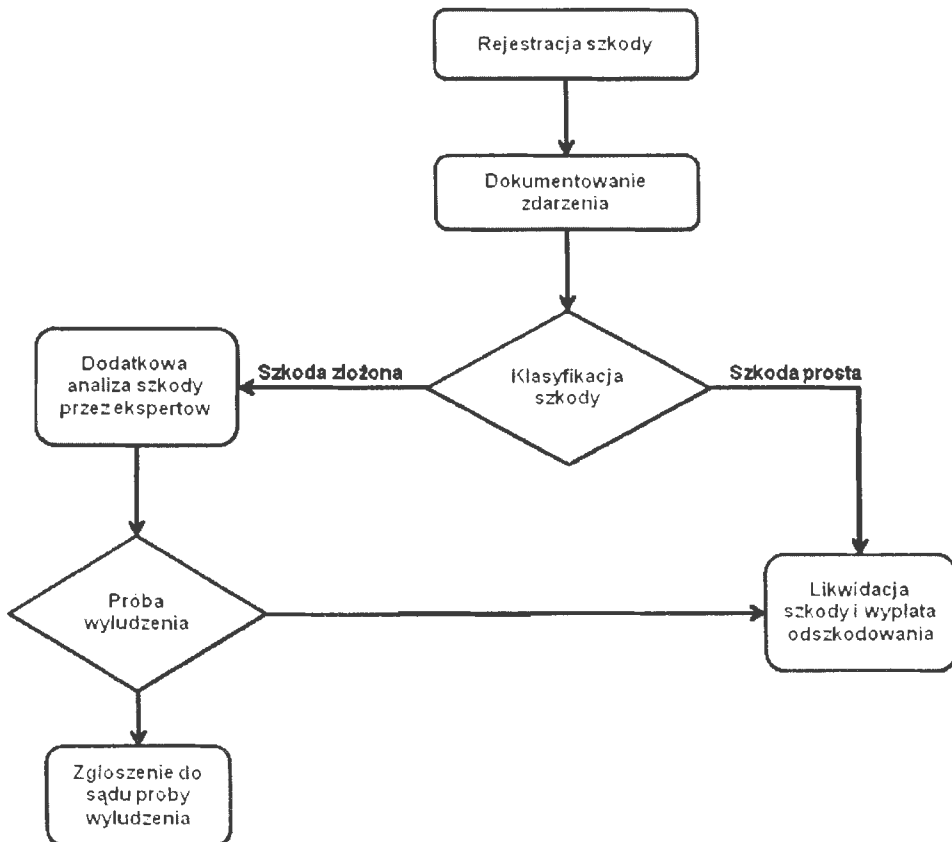
Nim jednak przejdę do opisu samego systemu przedstawię podstawowe informacje na temat wyłudzeń, likwidacji szkód oraz zbiorów rozmytych.

1. Rodzaje wyłudzeń komunikacyjnych

Firmy ubezpieczeniowe stykają się z coraz szerszą gamą sposobów jakimi próbuje się wyłudzić odszkodowanie.

Najczęściej spotykanymi są:

- fingowanie kradzieży drogich samochodów, które stanowią własność banków,
- pobieranie odszkodowań z kilku zakładów ubezpieczeniowych,



Rys. 1. Schemat procesu likwidacji szkody

- usiłowanie wyłudzeń za przedmioty, które nie znajdowały się w pojeździe czasie włamania do samochodu,
- umawianie się z osobą, która posiada AC na rzekome przyznanie się do spowodowania kolizji drogowej,
- nieuczciwe wyliczenia odszkodowania przez likwidatora,
- zawyżanie kosztów naprawy, przez fałszowanie faktur, zakupów części które nie uległy uszkodzeniu w wyniku zgłoszonej szkody.

Przedstawiona powyżej lista jest oczywiście niepełna i stale się powiększa. Na jej przykładzie widać jak potrzebna i trudna jest odpowiednia analiza zgłaszanych szkód na każdym etapie jej likwidacji.

2. Likwidacja szkód

Likwidacja szkody jest procesem wieloetapowym. Na każdym z tychże etapów może pojawić się niebezpieczeństwo nadużyć.

Sam proces możemy przedstawić w sposób pokazany na rys. 1.

W ogólnym przypadku dogłębnie analizuje się jedynie niektóre szkody. Może to prowadzić do pominięcia szkód w których występują ukryte elementy wyłudzenia.

Dlatego też warto by wiedze ekspertów zastosować do oceny nie tylko wybranych szkód, ale do wszystkich zgłoszeń. Niestety przy ogromnej ilości zgłaszanych codziennie szkód jest to zadanie bardzo trudne do wykonania. Nawet przez duży zespół ludzi. Warto więc spróbować ten proces choć w części zautomatyzować, a można to zrobić wykorzystując zbiory rozmyte.

3. Zbiory rozmyte

Język naturalny jest pełen nieprecyzyjnych, niedających się zdefiniować w sposób jednoznaczny wyrażeń. Na takie samo pytanie dotyczące tego samego obiektu (np. czy woda w basenie jest ciepła) możemy otrzymać zupełnie różne odpowiedzi (dla niektórych woda będzie zimna dla innych może być gorąca). Rozwiązanie problemu z matematycznym opisem takiej sytuacji opisał w swojej pracy z 1965 roku L. Zadeha.

Zadeh zaproponował by prócz określenia czy coś należy lub nie należy do danego zbioru (co możemy interpretować jako wartości prawda oraz fałsz) dopuścić nieskończenie wiele wartości pośrednich, czyli możliwość częściowego należenia do zbioru.

Obiekty matematyczne umożliwiające takie rozwiązanie nazywamy zbiorami rozmytymi (ang. fuzzy set).

Formalnie zaś zbiór rozmyty A w pewnej (niepustej) przestrzeni \mathbf{X} ($A \subseteq \mathbf{X}$),

zdefiniujemy jako zbiór uporządkowanych par:

$$A = \{(x, \mu_A(x)); x \in \mathbf{X}\}$$

gdzie $\mu_A : \mathbf{X} \rightarrow [0,1]$ jest tak zwaną funkcją przynależności zbioru rozmytego A .

Funkcja taka przypisuje każdemu elementowi $x \in \mathbf{X}$ wartość przynależności do zbioru rozmytego A , przy czym możemy wyróżnić trzy przypadki:

- $\mu_A(x) = 0$ oznacza brak przynależności elementu x do zbioru rozmytego A , tzn. $x \notin A$,
- $0 < \mu_A(x) < 1$ oznacza częściową przynależność elementu x do zbioru rozmytego A ,
- $\mu_A(x) = 1$ oznacza pełną przynależność do zbioru rozmytego A , tzn. $x \in A$.

4. Koncepcja systemu

System analizujący zgłoszone szkody ma na celu wsparcie ekspertów oraz osób zajmujących się likwidacją w ocenie prawdopodobieństwa wystąpienia wyłudzenia w zgłoszeniu.

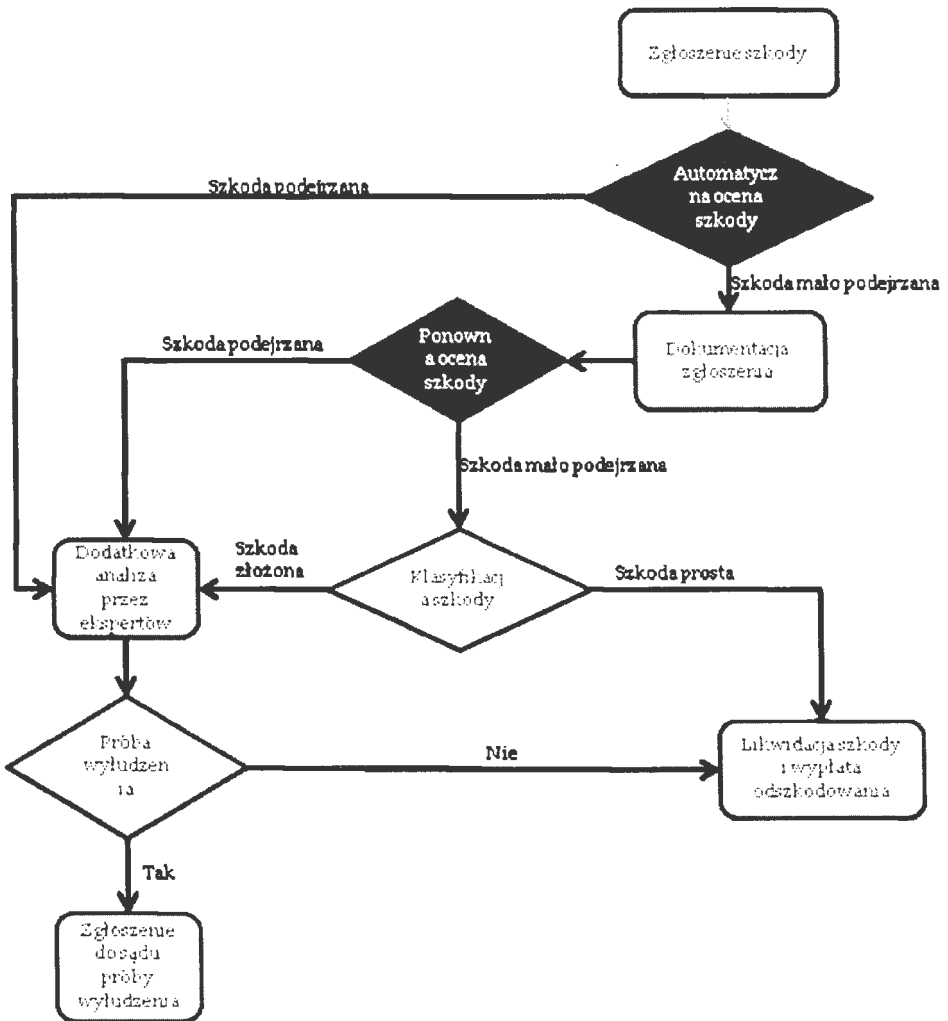
Aby móc to robić powinien mieć kilka kluczowych własności:

- powinien działać w czasie rzeczywistym, tj. analizować i oceniać szkodę w czasie aktualizacji dokumentacji jej dotyczącej,
- z możliwie szybko i dużym prawdopodobieństwem określać szkody jako niepodejrzane,
- posiadać łatwy w aktualizacji model oceny szkody
- pozwalać na wykorzystanie w nim wiedzy ekspertów.

Do tego celu idealnie nadają się zbiory rozmyte. Dzięki nim w dość prosty sposób można zapisać reguły na podstawie których eksperci „ręcznie” oceniają

szkodę. Dzięki temu na każdym etapie likwidacji szkody pozwala on na jego płynną analizę.

Po dodaniu systemu proces likwidacji szkody będzie wyglądał jak na rys. 2.



Rys. 2. Schemat likwidacji szkody z uwzględnieniem systemu automatycznej oceny

5. Analizowane przedziały

Jak wspomniałem wcześniej pożądanym jest by system mógł wykorzystać wiedzę ekspertów do oceny danej szkody. Informacje, które są przez nich używane do analizy, można zapisać w postaci reguł i przedziałów dla zbiorów rozmytych, a które po analizie podadzą liczbową informację na temat stopnia podejrzenia danej szkody. W zależności od tej informacji powinny być podejmowane odpowiednie działania.

Do analizy można wykorzystać informacje o:

- Miejscu wypadku (odległość od terenu zabudowanego),
- Szacowanej wysokości odszkodowania (kwoty małe, średnie, duże),
- Czasie pomiędzy datą zawarcia polisy a wypadkiem,
- Ilości szkód które były zgłoszone dla danego pojazdu,
- Porze dnia,
- Szkodowości osoby (liczba szkód zgłaszanych przez ubezpieczonego),
- Wieku kierującego,
- Złożoności szkody,
- Terminie zgłoszenia szkody (od razu po zdarzeniu, kilka dni później),
- Rodzaju pojazdu,
- Obecności policji.

Dla każdej z tych informacji budujemy zbiór rozmyty z odpowiednimi przedziałami, a następnie przy pomocy zdefiniowanych reguł oceniamy. Wykorzystując cechy zbiorów rozmytych do implikacji obszarów danych możemy łatwo połączyć reguły i obszary z kilku reguł w jedną. Dzięki temu uzyskamy jednoznaczną informację o ocenie szkody uwzględniającą wszystkie badane przez nas cechy.

Przykład

Dla lepszego zobrazowania działania systemu przeanalizujmy przykład dotyczący informacji o porze dnia wystąpienia szkody. Mamy szkodę która wystąpiła o godzinie 4, zbiór rozmyty.

Dodatkowo mamy dane reguły ekspertów które mówią, że:

- Jeżeli zdarzenie miało miejsce w nocy to jest bardzo podejrzane,
- Jeżeli zdarzenie miało miejsce rankiem to jest podejrzane.

Nanieśmy informację o godzinie zdarzenia na nasz zbiór:



Rys. 1. Przykład zbioru rozmytego z informacją na temat pory zdarzenia

Widzimy, iż linia przecina dwa obszary, a ich wartości odczytujemy na osi Y. Otrzymujemy odpowiednio wynik: $\text{noc} = 0.25$ oraz $\text{poranek} = 0.75$

Nakładamy teraz uzyskane informacje na zbiór rozmyty definiujący stopień podejrzania szkody:



Rys. 2. Zbiór rozmyty dotyczący stopnia podejrzania szkody

Na podstawie reguł i naniesienia wartości na przedziały stopnia podejrzania szkody, otrzymujemy szary obszar z którego obliczany środek ciężkości, który wynosi on 73. 21 czyli wynik należy odczytać jako **szkoda bardzo podejrzana = 0.66** i **szkoda podejrzana = 0.37**

Podsumowanie

Zaproponowany system może być niezwykle pomocny i efektywny w swoim działaniu. Jego dynamiczna struktura pozwala na ciągłe doskonalenie poprzez dodawanie kolejnych reguł i przedziałów. Pozwala on dość szybko wykryć ewentualną próbę wyludzenia i skierować dane zgłoszenia na odpowiednią ścieżkę likwidacji.

Literatura

- [1]. <http://www.piu.org.pl/> - Polska Izba Ubezpieczeń
- [2]. Jacek Łęski (2008): Systemy neuronowo-rozmyte. WNT, Warszawa.
- [3]. Dokumentacja wewnętrzna oprogramowania. Biuro Rozwoju Aplikacji Biznesowych, PZU S.A.

ISBN 9788389475220