PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1128 — 2006
Wpływ norm ostrożnościowych na ograniczanie ryzyka
w działalności banku spółdzielczego

Dariusz Wawrzyniak
Akademia Ekonomiczna we Wrocławiu

Józef Myrczek
Akademia Techniczno-Humanistyczna w Bielsku-Białej, BS Katowice

WYBRANE PROBLEMY SZACOWANIA POZIOMU RYZYKA INFORMATYCZNEGO W BANKOWOŚCI SPÓŁDZIELCZEJ

1. Wstęp

Ryzyko informatyczne związane jest z każdym przejawem zaawansowanej działalności gospodarczej, także działalności bankowej. Co więcej, złożoność technologiczna i funkcjonalna rozwiązań informatycznych stosowanych w bankowości, waga przetwarzanych w tych systemach informacji oraz wymogi prawa nadają problematyce zarządzania ryzykiem informatycznym znaczenie kluczowe dla funkcjonowania instytucji bankowych. Niezwykle istotnym aspektem procesu zarządzania tym rodzajem ryzyka jest możliwość ilościowego szacowania jego poziomu. Artykuł niniejszy jest próbą przedstawienia wybranej problematyki zastosowań metod ilościowych w zarządzaniu ryzykiem informatycznym w bankowości.

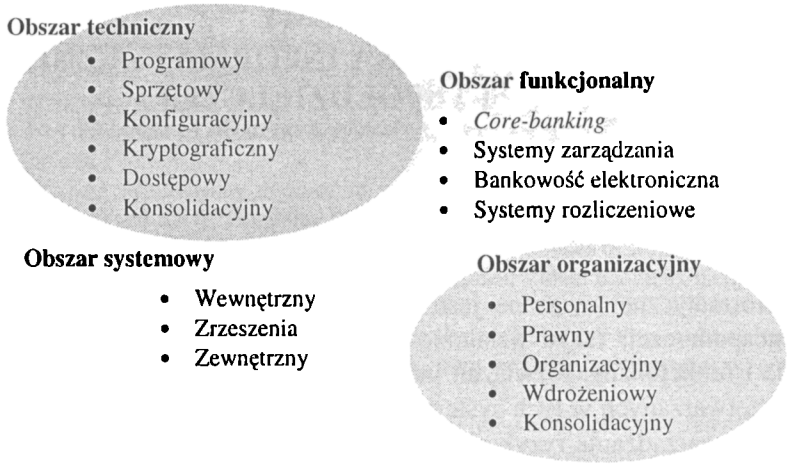
2. Obszary ryzyka w działalności bankowej

W klasycznym ujęciu na ryzyko działalności bankowej wpływa wiele rodzajów ryzyka: kredytowe, płynności, rynkowe, zmian systemowych oraz tworzące odrębną grupę typy niefinansowe określane jako ryzyko operacyjne. Rekomendacje Narodowego Banku Polskiego[1] oraz praktyka bankowa nakazują wyróżniać w ramach

[1] Rekomendacja D, dotycząca zarządzania ryzykami towarzyszącymi systemom informatycznym i telekomunikacyjnym używanym przez banki (2002) oraz Rekomendacja M, dotycząca zarządzania ryzykiem operacyjnym w bankach (2004).

ryzyka operacyjnego ryzyko organizacyjne, personalne, otoczenia i informatyczne. Za fundamentalny element ryzyka operacyjnego powszechnie uważane jest właśnie ryzyko informatyczne, definiowane przez Polską Normę PN-I-02000[2] jako możliwość, że konkretne zagrożenie wykorzysta konkretną podatność systemu przetwarzania danych. Zarządzanie ryzykiem informatycznym definiowane jest natomiast przez PN-ISO/IEC 17799[3] jako proces identyfikowania, kontrolowania i minimalizowania lub eliminowania ryzyka dotyczącego bezpieczeństwa przy zachowaniu akceptowalnego poziomu kosztów.

Ryzyko informatyczne we współczesnej bankowości powinno podlegać interdyscyplinarnej segmentacji obszarowej, ukierunkowanej na identyfikację odmiennych jego atrybutów, postrzeganych i analizowanych w różnych obszarach. Najważniejsze dla bankowości spółdzielczej obszary zarządzania ryzykiem przedstawiono na rys. 1.



Rys. 1. Obszary ryzyka informatycznego w bankowości spółdzielczej

Źródło: opracowanie własne.

Do obszarów tych zaliczamy:

- obszar techniczny, związany z wykorzystywanym sprzętem i oprogramowaniem oraz jego konfiguracją, problemami zastosowań współczesnych metod kryptograficznych, technicznymi aspektami zarządzania dostępem do zasobów systemowych, a także zagadnieniami technicznej integracji systemów w procesach konsolidacyjnych;

[2] PN-I-02000 – Technika informatyczna – Zabezpieczenia w systemach informatycznych – Terminologia, Polski Komitet Normalizacyjny, 1998.

[3] PN-ISO/IEC 17799 Technika Informatyczna. Praktyczne zasady zarządzania bezpieczeństwem informacji, Polski Komitet Normalizacyjny, 2003.

- obszar funkcjonalny, w ramach którego systemy informatyczne stosowane w instytucji bankowej podzielić można według ich funkcji na *core banking*, systemy wspomagające zarządzanie (finanse, logistyka, kadry i płace, CRM itp.), systemy bankowości internetowej, elektronicznej, systemy rozliczeniowe i inne;
- obszar systemowy, oddzielający systemy poszczególnych jednostek bankowych od systemów zrzeszenia oraz od systemów bankowości elektronicznej, do których dostęp posiadają także klienci banków – podział taki jest jednym ze specyficznych elementów charakterystycznych dla bankowości spółdzielczej;
- obszar organizacyjny, podkreślający konieczność zarządzania ryzykiem także w obszarach pozatechnicznych, dotyczących czynnika ludzkiego, strukturalnych i organizacyjnych, wdrożeniowych i konsolidacyjnych oraz w obszarze aspektów prawnych, wśród których przede wszystkim wyróżnić należy konieczność funkcjonowania w zgodzie z przepisami m.in. takich ustaw, jak:

– ustawa *Prawo bankowe*,
– ustawa o funkcjonowaniu banków spółdzielczych, ich zrzeszaniu się i bankach zrzeszających,
– ustawa o ochronie tajemnicy państwowej i służbowej,
– ustawa o ochronie danych osobowych,
– ustawa o ochronie informacji niejawnych,
– ustawa *Kodeks karny*,
– ustawa o elektronicznych instrumentach płatniczych,
– ustawa o świadczeniu usług drogą elektroniczną,
– ustawa o podpisie elektronicznym[4].

Na tle rozważań ukierunkowanych na specyfikę bankowości spółdzielczej na szczególną uwagę zasługuje podział obszaru systemowego na podobszary:

– wewnętrzny banku (spółdzielczego lub zrzeszającego),
– wewnętrzny zrzeszenia (obejmujący banki spółdzielcze i bank zrzeszający),
– zewnętrzny.

Podział taki stanowi o istotnej różnicy między bankowością spółdzielczą i komercyjną. Co więcej, obszar wewnętrzny zrzeszenia charakteryzuje się bardzo wysoką podatnością na różnego rodzaju zagrożenia bezpieczeństwa informatycznego. Systemy związane z obsługą zrzeszenia, będące istotnym źródłem ryzyka informatycznego, to m.in.:

– system rozliczeń wewnętrznych ze wspólnym „wyjściem" do KIR,
– system zarządzający kartami płatniczymi wydanymi w zrzeszeniu oraz bankomatami,
– system rozliczania dopłat do kredytów preferencyjnych z dopłatami ARiMR,
– system komunikacyjno-informacyjny.

[4] Szersze ujęcie prawnych aspektów bezpieczeństwa informatycznego w bankowości przedstawia m.in. A. Gospodarowicz [2005].

Podkreślenia wymaga bardzo istotny wpływ niektórych z wyżej wymienionych obszarów na współczesne problemy bankowości spółdzielczej w naszym kraju, procesy restrukturyzacyjne obserwowane w tym sektorze bankowym wymagają bowiem niezwykłej staranności, zwłaszcza w obszarze technicznym i organizacyjnym. Ten pierwszy wiąże się w omawianym kontekście z problemami integracji systemowej, zapewnienia poufności, integralności i dostępności danych, zachowania procedur bezpieczeństwa przy zmianach bądź rozszerzeniach systemów, monitorowania i dokumentowania tych zmian. Drugi natomiast to ogół zagadnień bezpośrednio związanych z działalnością czynnika ludzkiego – najsłabszego i najbardziej podatnego na zagrożenia elementu każdego systemu.

3. Ryzyko informatyczne w dokumentach Komitetu Bazylejskiego

Problematyka ryzyka informatycznego jest m.in. przedmiotem prac regulacyjnych Komitetu Bazylejskiego[5]. Fundamentem bieżących i przyszłych działań realizowanych w ramach zarządzania ryzykiem operacyjnym podmiotów systemu bankowego będą regulacje Nowej umowy kapitałowej oraz pochodne im dyrektywy Unii Europejskiej oraz rekomendacje Narodowego Banku Polskiego. Nowa umowa kapitałowa koncentruje się wokół ogółu problemów związanych z ryzykiem operacyjnym, definiowanym jako ryzyko strat w wyniku niewłaściwego lub błędnego działania procesu, ludzi i systemów lub wpływu wydarzeń wewnętrznych. W dokumentach komitetu znaleźć można także ramowy podział ryzyka operacyjnego na:

- ryzyko oszustwa ze strony pracowników,
- ryzyko oszustwa pochodzące z zewnątrz,
- ryzyko w zakresie zasad zatrudniania i bhp,
- ryzyko w zakresie zasad pracy z klientami i produktami,
- ryzyko szkód zasobów materialnych,
- ryzyko zakłócenia prowadzenia biznesu i niesprawności systemu,
- ryzyko zarządzania wykonywaniem zadań.

Komitet podjął również próbę nieco bardziej szczegółowego odniesienia się do omawianej problematyki w kontekście bankowości elektronicznej. Zasady zarządzania ryzykiem w bankowości elektronicznej [Risk Management Principles... 2003] to dokument będący swego rodzaju zbiorem ogólnych rekomendacji Komitetu, dotyczących tego obszaru działalności bankowej, który bezpośrednio wpływa na szeroko rozumiane bezpieczeństwo systemów bankowości elektronicznej. W dokumencie sformułowanych zostało między innymi 14 zasad zarządzania ryzykiem, które wyznaczają cele działań ukierunkowanych na zapewnienie optymalnego poziomu bezpieczeństwa konkretnych rozwiązań. Zasady te podzielono na trzy grupy:

[5] Bazylejski Komitet ds. Nadzoru Bankowego, http://www.bis.org/bcbs/. Tłumaczenia wybranych dokumentów znajdują się na stronie Narodowego Banku Polskiego.

1) kontrola ze strony rady i zarządu,
2) mechanizmy kontroli bezpieczeństwa,
3) zarządzanie ryzykiem prawnym i ryzykiem reputacji.

Wynikiem prac Komitetu Bazylejskiego jest także zbiór rekomendacji dotyczących szacowania wymogów kapitałowych związanych z ryzykiem operacyjnym. Spośród trzech propozycji – podejścia uproszczonego, standardowego oraz zaawansowanego – wśród banków komercyjnych (szczególnie tych zależnych od kapitału zagranicznego) najprawdopodobniej przyjmie się to ostatnie, bazujące na wewnętrznych modelach konstruowanych przez banki[6]. Banki spółdzielcze natomiast wybiorą raczej metody uproszczone lub standardowe, aczkolwiek nie jest wykluczone, że w najbliższej przyszłości zaistnieje konieczność konstruowania modeli bardziej zaawansowanych i – co niezwykle istotne – wspólnych dla całego sektora.

Problematyka szacowania wymogów kapitałowych może być swego rodzaju wprowadzeniem do zagadnień związanych ze szczegółową oceną i analizą poziomu ryzyka informatycznego na potrzeby zarządzania bezpieczeństwem. Wprawdzie istotne wydaje się zidentyfikowanie różnic między charakterem procedur analityczno-ewaluacyjnych w obu wspomnianych obszarach, niemniej wspólny mianownik w postaci ilościowego ujęcia poziomu ryzyka informatycznego te zagadnienia niewątpliwie łączy.

W dalszej części artykułu przedstawiono niektóre narzędzia ilościowych ocen ryzyka informatycznego związanego z wybranymi jego obszarami. Narzędzia te mogą być wykorzystywane jako samodzielne metody ukierunkowane na analizę wybranych obszarów systemu bądź stanowić składniki kompleksowych metodologii wspomagających procesy szacowania poziomu ryzyka informatycznego na potrzeby zarządzania bezpieczeństwem informatycznym w działalności bankowej.

4. Klasyfikacja metod szacowania ryzyka

Metody szacowania poziomu ryzyka informatycznego można podzielić według trzech podstawowych kryteriów: poziomu szczegółowości, typu modelu, na którym bazuje metoda, oraz rodzaju informacji wyjściowej.

W ramach podziału według kryterium poziomu szczegółowości wyróżnić można:
- metody ogólne (kompleksowe), które koncentrują się na ocenie relatywnie szerokiego spektrum zagadnień, z oceną ryzyka całych systemów włącznie,
- metody szczegółowe (cząstkowe), których celem jest ocena jedynie wybranych obszarów lub podobszarów systemu informatycznego.

[6] Nie jest wykluczone, że powszechne zastosowanie znajdą także inne rozwiązania metodyczne, promowane przez instytucje profesjonalnie zajmujące się audytem informatycznym.

W ramach kryterium związanego z typem modelu można wyróżnić:
- modele probabilistyczne, których podstawą są prawdopodobieństwa wystąpienia określonych zdarzeń,
- modele deterministyczne, w których występuje funkcyjna zależność między zmienną objaśnianą a zmiennymi objaśniającymi.

Według kryterium rodzaju informacji wyjściowej wyróżnia się natomiast:
- metody jakościowe, nie operujące na danych liczbowych, przedstawiające wyniki w postaci opisów i zaleceń,
- metody ilościowe, wykorzystujące dane liczbowe i prezentujące wyniki w postaci wskaźników.

Poniżej przedstawione zostały – w dość uproszczonym ujęciu – teoretyczne podstawy wybranych metod ilościowych wspomagających szacowanie poziomu ryzyka informatycznego. Są to metody szczegółowe, pozwalające na ilościowy opis wybranych obszarów funkcjonowania systemu informatycznego. Konstrukcja metody ogólnej, wykorzystującej np. przedstawione poniżej rozwiązania, musiałaby być ściśle skorelowana z wymogami informacyjnymi procesu zarządzania ryzykiem w danej instytucji bankowej. Innymi słowy, specyfiką syntetycznych wskaźników poziomu ryzyka informatycznego jest ich ścisły związek ze stawianymi przed nimi wymogami informacyjnymi, definiowanymi w ramach procesu zarządzania bezpieczeństwem. Zaproponowane metody mogą z powodzeniem wspomagać zarządzanie ryzykiem informatycznym w bankowości spółdzielczej.

5. Metody straty oczekiwanej

Pojęciem straty oczekiwanej posługuje się grupa metod ilościowych koncentrujących się wokół stochastycznego podejścia do omawianego zagadnienia. Strata oczekiwana jest wielkością charakterystyczną dla zdarzenia o negatywnym wpływie na bezpieczeństwo systemu informatycznego. Można zatem mówić o stracie oczekiwanej jako o finansowym odpowiedniku realizacji zagrożenia bezpieczeństwa systemu informatycznego. Wielkość straty SO w klasycznym ujęciu jest iloczynem prawdopodobieństwa wystąpienia danego zdarzenia $P(Z)$ oraz potencjalnej straty finansowej S, jaką to zdarzenie może spowodować.

$$SO = P(Z) \cdot S. \tag{1}$$

Postać zależności wiążącej prawdopodobieństwo ze stratą może być oczywiście inna. Przykładem takiego rozwiązania jest metoda Courtneya[7], w której stratę oczekiwaną wyznacza się następująco:

$$SO = \frac{1}{3} \cdot 10^{(k+c-3)}, \tag{2}$$

[7] Metoda zaproponowana w 1977 r. przez R. Courtneya [Caelli, Longley, Shain 1994, s. 92 i nast.].

gdzie k to wskaźnik straty finansowej, a c to wskaźnik częstości występowania zdarzenia. Wartości k i c wyznacza się za pomocą tabel (tab. 1, tab. 2).

Tabela 1. Strata finansowa wynikająca z realizacji zagrożenia

Koszt w przyjętych jednostkach	Wskaźnik k
10	1
100	2
1000	3
10000	4
100000	5
1000000	6

Tabela 2. Częstość występowania zagrożenia

Częstość	Wskaźnik c
Co 300 lat	1
Co 30 lat	2
Co 3 lata	3
Co 100 dni	4
Co 10 dni	5
Codziennie	6
10 razy dziennie	7

Przyjęte w metodzie prawdopodobieństwo wystąpienia (częstość występowania) może wynikać z analizy występowania danego zdarzenia w przeszłości bądź z arbitralnej decyzji eksperckiej. Wartość potencjalnej straty może mieć podobne źródła. Oczywiście wartości zaproponowane w tabelach mogą być modyfikowane na potrzeby konkretnych implementacji. Metoda najczęściej jest stosowana w celu skorelowania poszczególnych zagrożeń oraz wartości potencjalnych strat, pozwalającego na przypisanie tych zagrożeń do grup zgodnie z przyjętą przez bank koncepcją hierarchiczną. Jedną z takich koncepcji może być rekomendowana przez National Institute of Standards and Technology idea macierzowych analiz poziomu ryzyka informatycznego, zgodnie z którą poziom ryzyka ocenia się w podziale na trzy kategorie: niską, średnią i wysoką (por. [Risk Management Guide... 2001]). Oczywiście tak uproszczone podejście jest jedynie przykładem możliwości zastosowań omawianej grupy metod.

Metody bazujące na wartości straty oczekiwanej są nierzadko podstawą dalszych analiz ukierunkowanych na określanie poziomu zwrotu z inwestycji związanych z mechanizmami bezpieczeństwa systemu informatycznego. Tabela 3 przedstawia przykład implementacji tego typu analizy w postaci kolejnych etapów szacowania poziomu wewnętrznej stopy zwrotu z inwestycji.

Tabela 3. Etapy wyznaczania wewnętrznej stopy zwrotu z inwestycji

Etap	Wartość	Symbol	Sposób wyznaczenia
Etap 1	oczekiwana strata	OS	Prawdopodobieństwo realizacji zagrożenia × × wartość potencjalnej straty finansowej
Etap 2	korzyść z tytułu zastosowanych zabezpieczeń	K	OS – OS z zabezpieczeniami
Etap 3	wartość dodana	WD	K + nowe możliwości
Etap 4	zwrot z inwestycji	ROI	WD/koszty zabezpieczeń
Etap 5	wewnętrzna stopa zwrotu	IRR	$K_0 = \sum_{t=1}^{n} \frac{WD_t - K_t}{(1 + IRR)^t}$

Źródło: opracowanie własne na podstawie [Schechter 2004].

Podsumowując, metody oparte na koncepcji oczekiwanej wartości straty pozwalają na proste szacowanie poziomów ryzyka związanego z poszczególnymi zagrożeniami. Ich trywialność oraz stochastyczny charakter ograniczają jednak możliwości wykorzystywania ich w złożonych metodologiach wspomagających zarządzanie ryzykiem informatycznym. Dodatkowym utrudnieniem implementacji tego typu mechanizmów jest brak pełnych i wiarygodnych danych historycznych.

Nieco szerszym od przedstawionego powyżej podejścia zakresem merytorycznym charakteryzuje się koncepcja zaproponowana w ISO/TR 13569[8]. Opiera się ona na zestawieniu zagrożeń, podatności i kategorii ryzyka oraz wykorzystaniu analizy macierzowej. Merytoryczny zakres rekomendowanych analiz jest wprawdzie dość szeroki, niemniej prosta analiza macierzowa bazująca na trójstopniowej ocenie ryzyka (wysokie, średnie, niskie) również wydaje się być zbyt ogólna, aby jej zastosowanie w działalności bankowej mogło przynieść oczekiwane rezultaty. Raport ISO/TR 13569 powinien jednak stanowić podstawę do budowania indywidualnych rozwiązań analizy ryzyka informatycznego w bankowości.

Innym ciekawym podejściem do omawianej problematyki jest BITS Calculator. Metodyka ta oferuje możliwość wykorzystania dwóch koncepcji klasyfikowania ryzyka:

- normy ISO-17799,
- Nowej umowy kapitałowej.

Rozwiązanie takie pozwala na przypisanie każdego analizowanego kryterium ryzyka do którejś z kategorii w obu klasyfikacjach. Zestawienie kategorii uwzględnionych w metodyce przedstawia tab. 4.

[8] Raport techniczny ISO/TR 13569: 1997(E) Guidelines for the Management of IT Security.

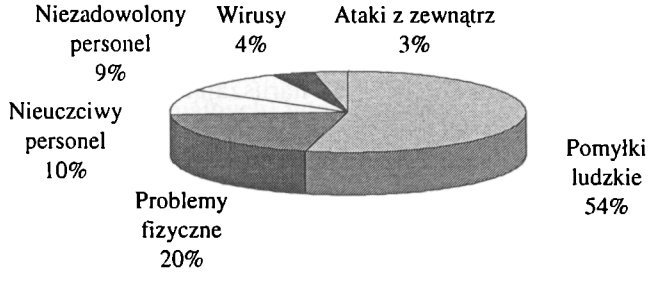
Tabela 4. Kategorie ryzyka w BITS Calculator

Według normy ISO 17799	Według Nowej umowy kapitałowej
Kontrola dostępu	Nadużycia wewnętrzne
Klasyfikacja zasobów sprzętowych i nadzór nad nimi	Nadużycia zewnętrzne
Zarządzanie ciągłością działania	Czynności pracowników, bezpieczeństwo i higiena pracy
Zarządzanie sferami telekomunikacji i eksploatacji	Klienci, produkty i czynności biznesowe
Zgodność z wymogami prawnymi	Uszkodzenia zasobów sprzętowych
Bezpieczeństwo organizacji	Awarie systemów i przerwy w działaniu
Bezpieczeństwo personelu	Zarządzanie wykonawstwem i procesami
Bezpieczeństwo fizyczne i środowiskowe	
Tworzenie systemów	

Źródło: [Pilawski 2005].

W metodyce wyróżnia się ponad sześćset zagrożeń pogrupowanych w wymienione powyżej kategorie. Macierz ocen operuje kombinacjami dwóch kryteriów: stopniem panowania organizacji nad daną przyczyną ryzyka i stopniem zagrożenia w przypadku utraty kontroli nad danym rodzajem ryzyka. Ilościowy aspekt BITS Calculator – podobnie jak w przypadku poprzednio omawianych koncepcji – jest nieskomplikowany, koncentruje się bowiem wokół wartości punktowych przypisywanych poszczególnym zagrożeniom.

W dalszej części artykułu przedstawiono wybór nieco bardziej złożonych metod pozwalających na szczegółowy opis ilościowy wybranych obszarów systemu informatycznego w banku. Najwięcej miejsca poświęcono metodom wspomagającym monitorowanie funkcjonowania systemów w kontekście działalności ich legalnych użytkowników. Statystyka pokazuje bowiem, jak znaczny udział (w ujęciu ilościowym) we wszystkich sytuacjach naruszających bezpieczeństwo systemu mają te związane z działalnością samych pracowników banków (rys. 2).



Rys. 2. Zagrożenia bezpieczeństwa systemów informatycznych w bankowości

Źródło: [Grzywacz 2003].

6. Metody wykrywania anomalii

Ryzyko informatyczne – jak już wspomniano – może być analizowane w różnych obszarach. Niewątpliwie jednym z najważniejszych we współczesnej bankowości jest wpływ działalności legalnych użytkowników systemów bankowości elektronicznej na bezpieczeństwo tych systemów. Wśród metod pozwalających na ilościowe ujęcie omawianego zagadnienia na uwagę zasługują przede wszystkim teoretyczne podstawy metod wykrywania anomalii w zachowaniach użytkowników systemów informatycznych.

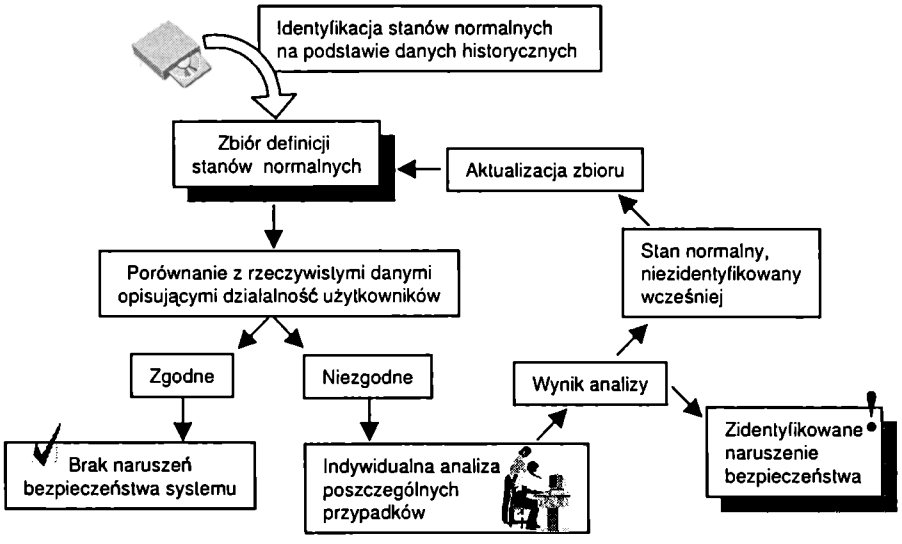
Literatura przedmiotu częściej niż pojęciem anomalii posługuje się pojęciem włamania, definiując je jako świadome działanie bądź ciąg współzależnych działań intruza mających na celu naruszenie bezpieczeństwa systemu[9]. Problematyka wykrywania włamań kojarzy się jednak głównie z nieautoryzowanym dostępem do zasobów informatycznych, podczas gdy nie mniej ważny – w bankowości nawet istotniejszy – jest problem działalności legalnych użytkowników. Anomalia, czyli odchylenie od normy, to w systemie informatycznym każde zdarzenie, które nie jest wynikiem standardowego zachowania się jego użytkownika. Innymi słowy, może być przejawem błędu operatorskiego (przypadkowego bądź świadomego), włamania intruza z zewnątrz lub – co szczególnie groźne dla systemów bankowych – celowego działania legalnego użytkownika systemu, ukierunkowanego na dokonanie nadużycia bądź naruszenie mechanizmów bezpieczeństwa systemu.

Metody wykrywania anomalii podzielić można na kilka typów, niemniej jednak wszystkie charakteryzują się określoną, schematyczną budową. Zasadę ich działania można w pewnym uproszczeniu przedstawić jako triadę:

określenie stanów normalnych → definicja stanów anormalnych → porównanie

Określenie stanów normalnych pozwala na stwierdzenie, co ma być uznane za stan nie będący efektem naruszenia bezpieczeństwa systemu, czyli wynik standardowego zachowania się użytkownika. Niejako pochodną określania stanów normalnych jest definicja stanów anormalnych. Można ją bowiem potraktować jako dopełnienie zbioru stanów normalnych. Porównanie zidentyfikowanych zbiorów stanów jest podstawą do stwierdzenia, czy w systemie miały miejsce zdarzenia anormalne, a więc naruszające jego bezpieczeństwo. Oczywiście nie każde zdarzenie zidentyfikowane jako anormalne musi być wynikiem ataku na system. Może się zdarzyć, że zastosowany algorytm wygenerował fałszywy alarm wynikający z nieoptymalnego zdefiniowania zbioru stanów normalnych. Reasumując, proces wykrywania anomalii można przedstawić schematycznie tak, jak pokazano na rys. 3.

[9] Za synonim terminu włamanie często przyjmuje się atak. Warto także nadmienić, że Polska Norma PN-I-02000 nie definiuje w ogóle włamania i anomalii, podając jedynie definicję ataku.



Rys. 3. Proces wykrywania anomalii

Źródło: opracowanie własne.

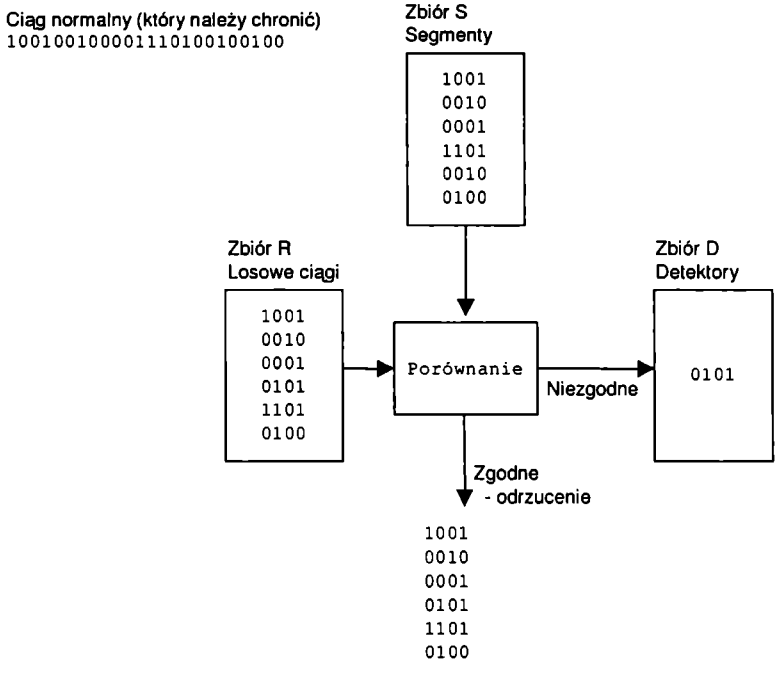
7. Algorytmy immunologiczne

Algorytmy te swoją budową i zasadami działania rzeczywiście nawiązują do mechanizmów funkcjonowania ludzkiej immunologii. Umożliwiają one identyfikację zdarzeń, które miały miejsce w systemie, a nie zostały wcześniej przyporządkowane do zbioru stanów normalnych.

Podstawowym zadaniem tych algorytmów jest stworzenie zbioru detektorów – ciągów reprezentujących normalne stany systemu[10]. W tym celu dokonuje się logicznego podziału danych reprezentujących stany normalne na segmenty o takiej samej długości. Należy oczywiście podkreślić, że jakość tych danych determinuje skuteczność algorytmu. Długość segmentów może być dowolna, tak samo jak ich postać (ciągi bitów, znaki ASCII). Na potrzeby teoretycznej analizy omawianej grupy algorytmów przyjęto, że segmentami będą n-wymiarowe ciągi bitów. Segmenty ze zbioru S (stanów normalnych) interpretowane są przez algorytm jako ciągi, które należy chronić. Innymi słowy, każdy ciąg nie należący do tego zbioru jest potencjalnym dowodem zaistnienia sytuacji naruszenia bezpieczeństwa systemu. Kolejnym etapem działania algorytmu jest wygenerowanie zbioru losowych ciągów bitów o takiej samej długości, jak ciągi ze zbioru S. Losowe ciągi tworzą zbiór R, którego elementy będą porównywane z elementami zbioru S. Operacja po-

[10] Algorytmy immunologiczne omówione zostały na przykładzie [Forrest, Perelson, Allen, Cherukuri 1997].

równania ma na celu stworzenie zbioru detektorów D, zawierającego tylko te ciągi z R, które nie znalazły swoich odpowiedników w S. Schemat tworzenia zbioru detektorów przedstawiono na rys. 4.



Rys. 4. Generowanie zbioru detektorów

Źródło: [Forrest, Perelson, Allen, Cherukuri 1997].

Praktyczna efektywność algorytmu wymusza jednak zaciemnienie przedstawionej powyżej, prostej i jasnej zasady działania. Wynika to przede wszystkim z konieczności:

- określenia optymalnej długości segmentów i losowych ciągów (kilku, czy nawet kilkunastobitowe ciągi byłyby bardzo nieefektywne),
- opracowania reguł częściowej zgodności ciągów,
- określenia docelowego, optymalnego rozmiaru zbioru detektorów.

Idealna zgodność dwóch ciągów o tej samej długości nie jest zjawiskiem trudnym do zidentyfikowania, chociaż zgodność taka – w przypadku ciągów kilkuset czy nawet kilkudziesięciobitowych – jest niezmiernie rzadko obserwowalna w praktyce. Może się zatem okazać, że praktyczne zastosowanie algorytmu wymaga określenia reguł zgodności częściowej. Jedną z najlepszych jest reguła zgodności kolejnych pozycji, która mówi, że ciągi x i y są zgodne, jeśli mają takie same bity na co najmniej r kolejnych pozycjach.

Przykładowo, na zero-jedynkowym alfabecie zidentyfikowano ciągi X i Y:

X 1000001010100101101110001001001000010
Y 0010011010001101010010001001001100000

Ciągi te są zgodne na maksymalnie 11 kolejnych pozycjach, zatem
ZGODNOŚĆ(X,Y) = PRAWDA dla $r \leq 11$, gdzie r jest wymaganą przez warunki porównania liczbą kolejnych pozycji, na których ciągi są zgodne.

Przy wyborze reguły zgodności częściowej warto przeanalizować prawdopodobieństwa zgodności dwóch ciągów na określonej liczbie pozycji. Określono je wzorem:

$$P_M \approx m^{-r}\left[(l-r)(m-1)/m+1\right], \tag{3}$$

gdzie: m – liczba znaków alfabetu, na bazie którego tworzone są ciągi,
l – liczba znaków w ciągu,
r – liczba kolejnych pozycji, na których muszą zgadzać się porównywane ciągi[11].

Tabela 5 przedstawia wartości P_M dla przykładowych argumentów.

Tabela 5. Prawdopodobieństwa zgodności ciągów

m	r	l	P_M
2	8	32	0,0502023
2	8	64	0,108697
2	8	128	0,2151
2	8	256	0,391316
2	16	32	0,000137329
2	16	64	0,000381437
2	16	128	0,000869474
2	16	256	0,00184483
128	8	32	$3{,}33067 \cdot 10^{-16}$
128	8	64	$7{,}77156 \cdot 10^{-16}$
128	8	128	$1{,}66533 \cdot 10^{-15}$
128	8	256	$3{,}44169 \cdot 10^{-15}$
128	16	32	~ 0,0
128	16	64	~ 0,0
128	16	128	~ 0,0
128	16	256	~ 0,0

Źródło: [Forrest, Perelson, Allen, Cherukuri 1997].

[11] Cytowani autorzy zaznaczają, że zależność jest prawdziwa jedynie dla $m^{-r} \ll 1$.

Zastosowanie algorytmu związane jest z pewnymi oczekiwaniami, przede wszystkim co do jego skuteczności, a więc prawdopodobieństwa wykrycia ciągu znaków, który nie został uprzednio sklasyfikowany jako normalny. Zakładając, że istnieją określone ciągi znaków, które należy chronić, można oszacować liczbę i długość detektorów wymaganych do zidentyfikowania ciągów będących efektem anormalnego zachowania się systemu.

Załóżmy, że:

N_{R0} – początkowa liczba detektorów (przed operacją generowania zbioru)[12],

N_R – liczba detektorów po operacji generowania zbioru detektorów,

N_S – liczba ciągów, które należy chronić,

P_M – prawdopodobieństwo zgodności pomiędzy dwoma losowymi ciągami,

f – prawdopodobieństwo, że losowy ciąg nie będzie zgodny z żadnym z N_S ciągów, które należy chronić = $(1 - P_M)^{N_S}$,

P_f – prawdopodobieństwo, że N_R detektorów nie wykryje anomalii.

Jeśli P_M jest małe, a N_S duże, to

$$f \approx e^{-P_M N_S} \tag{4}$$

oraz

$$N_R = N_{R_0} \cdot f, \tag{5}$$

$$P_f = (1 - P_M)^{N_R}. \tag{6}$$

Jeśli P_M jest małe, a N_R duże, to

$$P_f \approx e^{-P_M N_R}. \tag{7}$$

Zatem

$$N_R = N_{R_0} \cdot f = \frac{-\ln P_f}{P_M}. \tag{8}$$

Rozwiązawszy (3) i (4) ze względu na N_{R0}, otrzymujemy

$$N_{R_0} = \frac{-\ln P_f}{P_M \times (1 - P_M)^{N_S}}. \tag{9}$$

Formuła ta umożliwia oszacowanie początkowej liczby detektorów, wymaganej do wykrycia ciągu znaków, będącego efektem wystąpienia anomalii.

[12] Proces generowania zbioru detektorów może bazować na istniejącym już zbiorze.

Podsumowując rozważania nad algorytmem immunologicznym jako metodą wykrywania anomalii w bankowych systemach informatycznych, należy podkreślić, że:

- Algorytm jest elastyczny – istnieje możliwość wyboru docelowego, oczekiwanego prawdopodobieństwa wykrycia anomalii oraz oszacowania wymaganej liczby detektorów.
- N_R jest wielkością niezależną od N_S dla stałych P_M i P_f, co oznacza, że liczebność zbioru detektorów nie musi być funkcyjną zależnością liczby ciągów, które należy chronić. Dzięki takiej własności algorytm może efektywnie (ekonomicznie) chronić ogromne zbiory danych.

Algorytmy immunologiczne są istotnym narzędziem monitoringu, pozwalają bowiem wykrywać zdarzenia anormalne na bardzo wysokim poziomie szczegółowości. Ich optymalne zastosowanie wiąże się wprawdzie z koniecznością przeprowadzenia wielu wstępnych analiz wykorzystujących aparat matematyczny, niemniej ich użyteczność – głównie w obszarze analizy i pomiaru ryzyka informatycznego w systemach bankowych – jest nie do przecenienia.

8. Algorytmy oparte na procesach Markova

Jeśli w pewnym zbiorze stanów obiekt przechodzi z jednego stanu do innego z określonym prawdopodobieństwem, które nie zależy od stanu poprzedniego, lecz jedynie od tego, w którym obiekt znajduje się w danej chwili, to proces taki można nazwać procesem Markova[13].

Tworzenie profilu użytkownika może być oparte na analizie standardowych, normalnych w danej sytuacji zachowań, a więc stanów, w których użytkownik się znajduje, oraz ich zmian. Stanem nazywać będziemy uruchomioną aplikację, program, funkcję bądź wykonane polecenie systemowe. Analizując sekwencje stanów konieczne do wykonywania powierzonych pracownikowi obowiązków, można opracować ilościowy model przejść pomiędzy tymi stanami. Innymi słowy, dla każdego użytkownika systemu – pracownika banku – można zdefiniować profil zachowań w postaci macierzy prawdopodobieństw zmian stanów. Dokładne określenie takiego profilu stanowi doskonałą podstawę do późniejszych działań monitorujących. Profile zachowań pozwalają na identyfikację zagrożeń bankowych systemów informatycznych, których źródłami są celowe lub niecelowe działania pracowników banku, w szczególności różnego rodzaju nadużycia.

Do opisu profilu użytkownika wykorzystać można macierz przejść, przedstawiającą prawdopodobieństwa przejść pomiędzy stanami (tab. 6).

[13] Szczegółowy opis teoretycznych podstaw tych procesów oraz ich praktycznych implikacji – głównie w sferze podejmowania strategicznych decyzji – znaleźć można m.in. w pracy M. Putermana [1994], a także w wydawnictwach polskojęzycznych (por. [Plucińska, Pluciński 2005; Wentzell 1980]).

Tabela 6. Macierz **M** prawdopodobieństw przejść pomiędzy stanami

	Stan 1	Stan 2	Stan 3	Stan 4	...	Stan *n*	
Stan 1	p_{11}	p_{12}	p_{13}	p_{14}		p_{1n}	$\sum_{i=1}^{n} p_{1i} = 1$
Stan 2	p_{21}	p_{22}	p_{23}	p_{24}		p_{2n}	$\sum_{i=1}^{n} p_{2i} = 1$
Stan 3	p_{31}	p_{32}	p_{33}	p_{34}		p_{3n}	$\sum_{i=1}^{n} p_{3i} = 1$
Stan 4	p_{41}	p_{42}	p_{43}	p_{44}		p_{4n}	$\sum_{i=1}^{n} p_{4i} = 1$
...							
Stan *n*	p_{n1}	p_{n2}	p_{n3}	p_{n4}		p_{nn}	$\sum_{i=1}^{n} p_{ni} = 1$

Źródło: opracowanie własne.

Poszczególne elementy macierzy przedstawiają postulowane (modelowe) prawdopodobieństwa przejść pomiędzy stanami dla określonego użytkownika. Macierz prawdopodobieństw przejść może być wyznaczana *a priori* przez zespół ekspertów bądź może być wynikiem obserwacji działającego systemu. Wykorzystanie macierzy w metodzie wykrywania anomalii polega na porównaniu jej z rzeczywistymi obserwacjami. Aby porównanie było możliwe, obserwacje rzeczywiste przedstawić należy także w postaci macierzy, której elementem o_{ij} będzie wartość odpowiadająca procentowemu udziałowi zaobserwowanych przejść ze stanu i do stanu j (k_{ij}) w stosunku do sumy przejść ze stanu i do innych stanów (S). Innymi słowy, będzie to macierz rzeczywistej częstości przejść (tab. 7).

Tabela 7. Macierz **O** obserwacji

	Stan 1	Stan 2	Stan 3	Stan 4	...	Stan *n*	
Stan 1	$o_{11}=k_{11}/S$	o_{12}	o_{13}	o_{14}		o_{1n}	$S = \sum_{i=1}^{n} k_{1i}$
Stan 2	o_{21}	o_{22}	o_{23}	o_{24}		o_{2n}	
Stan 3	o_{31}	o_{32}	o_{33}	o_{34}		o_{3n}	
Stan 4	o_{41}	o_{42}	o_{43}	o_{44}		o_{4n}	
...							
Stan *n*	o_{n1}	o_{n2}	o_{n3}	o_{n4}		o_{nn}	

Źródło: opracowanie własne.

Tak jak poprzednio, powstaje problem optymalnych relacji porównawczych. Możliwości jest tutaj wiele, jednak wszystkie sprowadzają się do wyznaczenia swego rodzaju odległości między macierzami **M** i **O**, odzwierciedlającej ich niezgodność. Odległość taką traktować można jako wskaźnik charakteryzujący rozbieżność między postulowanym modelem zachowań a rzeczywistością. Postać re-

lacji porównawczej zależy od przyjętych kryteriów analizy, można bowiem porównywać nie tylko przejścia jednokrokowe, lecz także wielokrokowe.

Dobrym rozwiązaniem porównawczym wydaje się być jedno z najprostszych – unormowany wskaźnik rozbieżności dla pojedynczych kroków R^1.

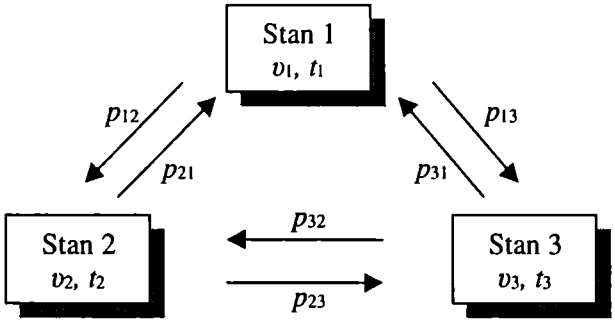
$$R^1 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|o_{ij} - p_{ij}\right|}{2n}, \tag{10}$$

gdzie n to wymiar macierzy **M** i **O**, a o i p – elementy tych macierzy.

Uzupełnieniem tak określonego wskaźnika[14] może być wskaźnik dla n kroków, a syntetyczną wartością obrazującą rozbieżność macierzy średnia ważona obu wskaźników. Określona przez ekspertów średnia ważona wartości R^1 dla wszystkich monitorowanych użytkowników powinna być podstawą przeciwdziałania występowaniu anomalii (działań minimalizujących ryzyko informatyczne, czyli podwyższających poziom bezpieczeństwa systemu informatycznego). Jej interpretacja zależy oczywiście od przyjętych wcześniej założeń. W połączeniu z innymi zaobserwowanymi parametrami systemu, powinna być podstawą do konstruowania syntetycznych mierników obrazujących poziom (lub zmiany poziomu) ryzyka informatycznego w banku. Przykładowo, postulowaną macierz przejść można uzupełnić o dodatkowe elementy, takie jak:

- częstotliwość przebywania w określonych stanach v,
- długość przebywania w tych stanach – średni czas pracy t,

tworząc rozszerzony graf stanów (rys. 5).



Rys. 5. Rozszerzony graf stanów

Źródło: opracowanie własne.

[14] Więcej wskaźników pozwalających na praktyczne zastosowanie metody zaproponowano w [Wawrzyniak 2002].

Oczywiście wielkości p, v i t porównywane są osobno, a końcowa wartość rozbieżności pomiędzy modelem i obserwacjami zależy od eksperckiej konstrukcji wskaźnika uniwersalnego.

9. Modele przewidywania zagrożeń

Modele tego typu w sposób istotny wspomagają zarządzanie ryzykiem informatycznym. Należą do grupy metod probabilistycznych i powinny stanowić bardzo ważny składnik każdego rozwiązania kompleksowego. Przykładem takiego modelu jest model przewidywania zagrożeń, w którym założono seryjne wykorzystywanie przez intruza (czyli np. pracownika banku) nieznanej administratorom podatności systemu. Dodatkowym założeniem jest istnienie uczącego się systemu wykrywania anomalii (por. [Schechter 2004]).

Zmienne modelu to:

L – przychód intruza uzyskany z pojedynczego ataku,

P_C – prawdopodobieństwo, że intruz zostanie złapany,

F – koszt intruza w przypadku przyłapania,

P_D – prawdopodobieństwo, że wykorzystanie podatności przyniesie skutek w postaci jej identyfikacji oraz zabezpieczenia przez administratora,

P_F – prawdopodobieństwo nieudanego ataku.

Oczekiwany przychód intruza z i-tego ataku wynosi (zakładając, że podatność nie została jeszcze zidentyfikowana i zabezpieczona):

$$Z_i = (1 - P_F)L - P_C F. \tag{11}$$

Oczekiwany przychód intruza z ataku $i+1$ wynosi:

$$Z_{i+1} = [(1 - P_F)L - P_C F](1 - P_D). \tag{12}$$

Oczekiwany przychód intruza z ataków i, $i + 1$, ..., n wynosi:

$$\begin{aligned} Z_{i \to n} = {} & [(1 - P_F)L - P_C F] + [(1 - P_F)L - P_C F](1 - P_D) + [(1 - P_F)L - \\ & - P_C F](1 - P_D)^2 + \ldots + [(1 - P_F)L - P_C F](1 - P_D)^n. \end{aligned} \tag{13}$$

Wiadomo, że $\sum_{i=0}^{\infty} x^i = \frac{1}{1 - x}$ dla $0 < x < 1$.

Jeśli $n \to \infty$, to:

$$Z_{i \to \infty} = \frac{(1 - P_F)L - P_C F}{P_D}. \tag{14}$$

Przedstawiona powyżej prosta koncepcja pozwala na ilościowe oszacowanie zarówno ryzyka związanego z określonymi zdarzeniami naruszającymi bezpieczeństwo systemu, jak i wpływu, jaki ma zwiększanie jakości systemu wykrywania anomalii na to ryzyko, widać bowiem wyraźnie, że oczekiwany przychód intruza jest odwrotnie proporcjonalny do skuteczności systemu wykrywania anomalii.

10. Podsumowanie

W referacie przedstawiono część zagadnień związanych z zastosowaniem metod ilościowych w zarządzaniu ryzykiem informatycznym systemów bankowych. Skoncentrowano się na problematyce możliwości zastosowań tych metod w kontekście monitoringu działalności legalnych użytkowników systemów informatycznych, ponieważ to właśnie pracownicy banków stanowią najpoważniejsze obecnie zagrożenie bezpieczeństwa systemów bankowych. W pracy nie dokonano kompletnego przeglądu metod ilościowych przydatnych w zarządzaniu ryzykiem informatycznym. Nie poruszono chociażby problemów dotyczących algorytmów bayesowskich czy algorytmów opartych na sieciach neuronowych, a także licznych modeli regresyjnych.

Obecnie za wcześnie jest na sformułowanie propozycji rozwiązania kompleksowego, co jednak może być usprawiedliwione tym, że tego typu rozwiązania muszą spełniać specyficzne wymagania konkretnych systemów i polityk bezpieczeństwa zaimplementowanych w instytucjach bankowych. Podkreślić także należy, że współczesne problemy bankowości spółdzielczej w naszym kraju zdecydowanie wymagają stosowania rozwiązań wspomagających proces zarządzania ryzykiem informatycznym. Ze względu na specyfikę ryzyka operacyjnego występującego w sektorze spółdzielczym jest to jednak trudniejsze niż w przypadku banków komercyjnych.

Literatura

Caelli W., Longley D., Shain M., *Information Security Handbook*, Macmillan Press Ltd. 1994.

Forrest S., Perelson A., Allen L., Cherukuri R., *Self-Nonself. Discrimination in a Computer*, IEEE Symposium on Security and Privacy, 1997.

Gospodarowicz A. (red.), *Bankowość elektroniczna*, PWE, Warszawa 2005.

Grzywacz J. (red.), *Bezpieczeństwo systemów informatycznych w bankach w Polsce*, SGH, Warszawa 2003.

Pilawski B., *Metody analizy i oceny ryzyka informatycznego – wnioski ze stosowania w praktyce bankowej*, [w:] *Ryzyko informatyczne w działalności bankowej*, IX Forum Bankowości Elektronicznej, Wydawnictwo Centrum Promocji Informatyki, Warszawa 2005.

Plucińska A., Pluciński E., *Probabilistyka. Rachunek prawdopodobieństwa. Statystyka matematyczna. Procesy stochastyczne*, Wydawnictwo Naukowo-Techniczne, Warszawa 2005.

Puterman M., *Markov Decision Processes*, John Wiley & Sons 1994.
Risk Management Principles for Electronic Banking, Basel Committee on Banking Supervision, 2003.
Risk Management Guide for Information Technology Systems, NIST Special Publication 800-30, 2001.
Schechter S.E., *Computer Security Strength & Risk: A Quantitative Approach*, Harvard University 2004.
Wawrzyniak D., *Zarządzanie bezpieczeństwem systemów informatycznych w bankowości*, Wydawnictwo Zarządzanie i Finanse, Warszawa 2002.

CHOSEN ASPECTS OF INFORMATION TECHNOLOGY RISK ASSESSMENT IN CO-OPERATIVE BANKING

Summary

The article presents chosen aspects of information technology in co-operative banking. The main areas of the IT risk were pointed out. Chosen Basel Committee documents were also presented as well as their influence on the risk management process. The necessity of using the quantitative methods was emphasized. The authors presented the general classification of the methods and described some of them pointing out the possibilities for taking the advantage of the methods in co-operative banking. Very important part of the article deals with behaviour of IT systems' users and its influence on the systems security. The authors described selected methods allowing monitoring of users' activities. The main focus was set on the immunological methods and the methods based on the Markov processes. Some easy mathematical models were also presented.