

PRACE NAUKOWE

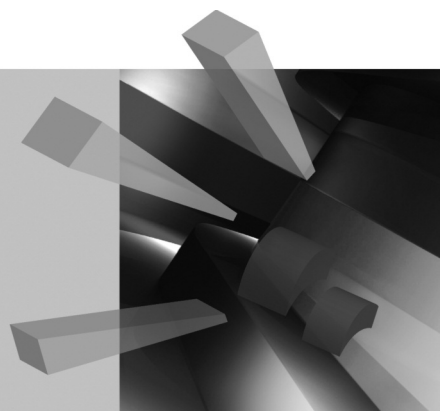
Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

236

Badania marketingowe – metody, nowe podejścia i konteksty badawcze



pod redakcją

Krystyny Mazurek-Łopacińskiej

Magdaleny Sobocińskiej



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2011

Recenzenci: Anna Dąbrowska, Lechosław Garbarski, Józef Garczarczyk

Redaktor Wydawnictwa: Agnieszka Flasińska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna na stronie www.ibuk.pl

Streszczenia opublikowanych artykułów są dostępne w międzynarodowej bazie danych The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl> oraz w The Central and Eastern European Online Library www.ceeol.com a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się na stronie internetowej Wydawnictwa www.wydawnictwo.ue.wroc.pl

Kopiowanie i powielanie w jakiegokolwiek formie wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2011

ISSN 1899-3192

ISBN 978-83-7695-248-2

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp.....	9
------------	---

Część 1. Rozwój badań marketingowych – kierunki, koncepcje, wymiary

Krystyna Mazurek-Łopacińska, Magdalena Sobocińska: Rozwój badań marketingowych – w kierunku nowych podejść i kontekstów badawczych związanych z funkcjonowaniem przedsiębiorstwa.....	13
Dariusz Oczachowski: Kultura kognitywna organizacji a prowadzenie badań marketingowych	23
Tomasz Heryszek: Myślenie kognitywne czy afektywne? Między danymi twardymi a przeczuciem – dylematy współczesnego przedsiębiorcy	33
Jan W. Wiktor: Koncepcja i sposób pomiaru internacjonalizacji przedsiębiorstwa w świetle doświadczeń projektu „Strategie marketingowe przedsiębiorstw na rynkach międzynarodowych”	42
Aleksandra Nizielska: Dystans kulturowy w procesie internacjonalizacji przedsiębiorstw – metodyczny aspekt badań.....	51
Magdalena Soboń, Iga Rudawska, Sylwia Bąkowska: Metodyka badania zachowań konsumenckich w międzynarodowej sieci badawczej.....	61
Bogusław Bembenek: Rola wywiadu gospodarczego w zarządzaniu wiedzą w klastrze	71

Część 2. Nowe podejścia badawcze wynikające z rozwoju technologii informacyjnych i komunikacyjnych

Mariusz Kuziak: Wyzwania wobec badań użytkowników Internetu.....	83
Karol Łopaciński: Model tematycznej hurtowni danych na potrzeby badania przebiegu i efektów kampanii e-mailingowych.....	95
Magdalena Jaciow: Fora internetowe jako źródło informacji – możliwości i ograniczenia wykorzystania w badaniach zachowań nabywczych	113
Beata Kolny: Wykorzystanie netnografii do badania usług zagospodarowujących czas wolny	123
Radosław Szulc, Piotr Ciszewski: Wykorzystanie technologii łączności bezprzewodowej w badaniach marketingowych na rynku produktów <i>beauty care</i>	132

Część 3. Metody ilościowe – nowe podejścia i zastosowania oraz triangulacja metod

Adam Sagan: Asymetryczne metody wielowymiarowe w badaniach marketingowych	145
Mariusz Łapczyński: Łączenie metod i narzędzi w budowie modeli predykcyjnych.....	155
Józef Garczarczyk, Robert Skikiewicz: Zastosowanie metody grupowania dwustopniowego w segmentacji klientów indywidualnych na rynku usług bankowych.....	164
Anna Bryja: Zmienne jakościowe w segmentacji rynku: miary powiązań a wyniki grupowania.....	175
Bartłomiej Jefmański: Nowe podejście w pomiarze opinii respondentów z zastosowaniem skal porządkowych i elementów teorii zbiorów rozmytych – charakterystyka wybranych aspektów metodologicznych.....	184
Grzegorz Maciejewski: Wykorzystanie analizy czynnikowej w badaniach konsumenckiego ryzyka	192
Paweł Chlipała: Zastosowanie eksperymentu w badaniach społecznie odpowiedzialnej konsumpcji – refleksje nad metodą, wyniki	203

Część 4. Badania jakościowe i ich wykorzystanie w rozwiązywaniu problemów badawczych i decyzyjnych

Marcin Komor: Znaczenie i rozwój metod jakościowych w badaniach empirycznych w marketingu.....	215
Sylwia Wrona: Dobór próby w jakościowych badaniach marketingowych – problemy prawidłowej selekcji i rekrutacji uczestników	225
Mateusz Rak, Joanna Nogiec: Wykorzystanie wyników badań jakościowych do identyfikacji populacji w badaniach ilościowych.....	234
Iwona Olejnik: Metoda obserwacji – zastosowania w badaniach marketingowych.....	242
Zbigniew Piskorz: Rozpoznawanie okazji przedsiębiorczych – rezultaty badań jakościowych.....	250
Zbigniew Spyra: Zastosowanie podejścia etnograficznego we współczesnych badaniach marketingowych w sferze kultury	260
Wanda Patrzalek: Przebieg i zaburzenia procesów wymiany informacji oraz komunikacji wewnątrz gospodarstwa domowego w ogólnopolskich badaniach fokusowych	271

Joanna Wardzała-Kordyś: Wiedza i opinie o procesach reklamacji produktów wadliwych w kontekście badań gospodarstw domowych	280
Agnieszka Dejnaka: Komunikacja pomiędzy członkami gospodarstwa domowego przy użyciu nowoczesnych narzędzi wymiany informacji	290
Jolanta Tkaczyk: Rola opowieści w badaniach marketingowych	301

Summaries

Part 1. Development of marketing research – trends, concepts, dimensions

Krystyna Mazurek-Łopacińska, Magdalena Sobocińska: Development of marketing research – towards new approaches and contexts of research related to the functioning of a company	22
Dariusz Oczachowski: Cognitive culture of an organization and conducting marketing research	32
Tomasz Heryszek: Cognitive or affective thinking? Between hard data and foreboding – modern business dilemmas	41
Jan W. Wiktor: The concept and measurement method of the companies' internationalization in the light of the research project "Marketing strategies of companies on international markets"	50
Aleksandra Nizielska: Cultural distance in the process of companies' internationalization – methodological aspect of research	60
Magdalena Soboń, Iga Rudawska, Sylwia Bąkowska: Consumer behaviour research methods in the International Research Network	70
Bogusław Bembenek: The role of economic intelligence in knowledge management of cluster	80

Part 2. New research approaches arising from the development of information and communication technologies

Mariusz Kuziak: Challenges to Internet audience measurement	94
Karol Łopaciński: Model of data mart prepared due to research regarding proceeding of e-mail campaigns and its effects	112
Magdalena Jaciow: Online forums as a source of information – possibilities and limitations of use in purchasing behaviour research	122
Beata Kolny: The application of netnographic surveys to research on leisure time services	131
Radosław Szulc, Piotr Ciszewski: The implementation of wireless technology in marketing research within the beauty industry	142

Part 3. Quantitative methods – new approaches, applications and triangulation methods

Adam Sagan: Asymmetric multivariate methods in marketing research	154
Mariusz Łapczyński: Combining methods and tools in building predictive models.....	163
Józef Garczarczyk, Robert Skikiewicz: Applying a Two Step Cluster method in the segmentation of individual customers for the banking services market	174
Anna Bryja: Qualitative variables in market segmentation: Similarity coefficients and clustering results.....	183
Bartłomiej Jefmański: A new approach in respondents' opinion measurement using ordinal scales and elements of fuzzy sets theory – characteristics of selected methodological aspects	191
Grzegorz Maciejewski: The use of factor analysis in consumer risk research	202
Paweł Chlipała: Using an experiment in the research of socially responsible consumption – reflections on the method, the results.....	212

Part 4. Qualitative research and its use in problems solving research and decision

Marcin Komor: The importance and development of quality methods in empirical research in marketing	224
Sylvia Wrona: Selection of a sample in qualitative marketing research – issues related to the accurate selection and recruitment of participants.....	233
Mateusz Rak, Joanna Nogiec: Using the results of qualitative research to the identification of the population in quantitative research.....	241
Iwona Olejnik: The method of observation – application in marketing research	249
Zbigniew Piskorz: Entrepreneurial opportunity recognition – results of qualitative research	259
Zbigniew Spyra: Application of ethnographic approach to contemporary marketing research in culture.....	270
Wanda Patrzalek: Proceeding and abnormal processes of information exchange and communication within the household in nationwide focus studies	279
Joanna Wardzała-Kordyś: Knowledge and opinions about complaints to the defective products in the context of household surveys	289
Agnieszka Dejnaka: Communication between members of the household by using modern information exchange tools.....	300
Jolanta Tkaczyk: The role of storytelling in marketing research	310

Karol Łopaciński

Uniwersytet Ekonomiczny we Wrocławiu

MODEL TEMATYCZNEJ HURTOWNI DANYCH NA POTRZEBY BADANIA PRZEBIEGU I EFEKTÓW KAMPANII E-MAILINGOWYCH

Streszczenie: W artykule zaprezentowane zostały projekty modelu pojęciowego i logicznego tematycznej hurtowni danych mającej zastosowanie w badaniu skuteczności kampanii e-mailingowych przedsiębiorstwa. Zadaniem hurtowni będzie gromadzenie, ujednolicanie i organizowanie danych na temat przebiegu i efektów kampanii realizowanych w różnych okresach. W ramach projektu modelu pojęciowego scharakteryzowano modelowany obszar biznesowy, użytkowników docelowych projektowanego narzędzia i źródła danych zasilające hurtownię, a także określono podstawowe fakty i wymiary hurtowni. Konstruując model logiczny, dokonano wyboru jego struktury, zaprojektowano tabelę faktów oraz tabele wymiarów wraz z ich atrybutami i elementami atrybutów. Ponadto zdefiniowano relacje łączące poszczególne tabele. Prototyp hurtowni tematycznej będzie stanowił podstawę przeprowadzania analiz wielowymiarowych przy użyciu narzędzi typu OLAP, których zadaniem będzie określanie skuteczności prowadzonych przez przedsiębiorstwo działań promocyjnych w Internecie.

Słowa kluczowe: promocja internetowa, kampanie e-mailingowe, pomiar skuteczności promocji w Internecie, tematyczna hurtownia danych, model pojęciowy, model logiczny.

1. Wstępne założenia konstrukcji modelu

Kampanie promocyjne w sieci mogą być prowadzone z użyciem różnorodnych narzędzi opartych na technologiach internetowych. Zakres możliwych do zastosowania rozwiązań jest obecnie bardzo duży. Wśród instrumentów wykorzystywanych przez organizacje do osiągania swoich celów promocyjnych w Internecie należy wymienić przede wszystkim firmową stronę WWW, tekstową i multimedialną reklamę odsłoną na portalach i serwisach internetowych, promocję w wyszukiwarkach, promocję za pośrednictwem poczty elektronicznej, internetowe *public relations*, promocję na portalach społecznościowych oraz marketing wirusowy.

W niniejszym artykule, którego istotą jest przedstawienie koncepcji modelu szacowania skuteczności promocji internetowej opartego na technologii Business Intelligence, autor skoncentrował się na propozycji wykorzystania możliwości wspomnianego modelu w kampaniach promocyjnych realizowanych z wykorzystaniem poczty elektronicznej (kampanii e-mailingowych).

Dane opisujące przebieg tego typu kampanii mogą być bardzo cennym materiałem w badaniu skuteczności działań promocyjnych przedsiębiorstwa w Internecie. Problem polega na tym, że dane te są zazwyczaj rozproszone i pochodzą z różnych źródeł, a także mają odmienne typy i formaty. Warunkiem ich przydatności jest zaprojektowanie właściwej bazy danych, wewnątrz której zasoby informacyjne dotyczące kolejnych kampanii będą sukcesywnie zapisywane i gromadzone oraz porządkowane i ujednolicane. Zorganizowane w ten sposób mogą być następnie odpowiednio przetwarzane i stanowić podstawę badań skuteczności danych kampanii z wykorzystaniem narzędzi analitycznych. Wyniki analiz w formie zwięzłych i kompleksowych zestawień mogą być następnie udostępniane za pośrednictwem narzędzi raportujących osobom zgłaszającym zapotrzebowanie na tego typu informacje.

Prowadzenie kampanii e-mailingowych należy uznać za bardzo wąski obszar działalności biznesowej przedsiębiorstwa. Dlatego też zasadne wydaje się, aby dane o szczegółowym przebiegu wspomnianych kampanii w czasie były gromadzone w tematycznych hurtowniach danych (*data mart*). Hurtownie tematyczne często nazywane są również hurtowniami OLAP (*OnLine Analytical Processing*), ponieważ celami ich budowy są organizacja i przetwarzanie danych *stricto* na potrzeby analiz i raportów OLAP.

Konkretyzując, w ramach artykułu skonstruowane zostaną modele pojęciowy i logiczny tematycznej hurtowni danych mającej zastosowanie w badaniu skuteczności kampanii e-mailingowych. Jej zadaniem będzie gromadzenie, ujednolicanie i organizowanie szczegółowych danych na temat przebiegu i efektów realizowanych w różnych okresach czasu kampanii promocyjnych. Prototyp hurtowni tematycznej będzie stanowił podstawę do przeprowadzania analiz wielowymiarowych z użyciem narzędzi typu OLAP.

Na potrzeby artykułu przeprowadzone zostały prace programistyczne, których rezultatem jest implementacja skonstruowanego modelu logicznego w rzeczywistej tematycznej hurtowni danych. Platformą bazodanową dla zaprojektowanej hurtowni danych jest Microsoft SQL Server 2008 R2. Hurtownię zasilono danymi źródłowymi, które zostały wygenerowane na bazie symulacji. Zgromadzone w ten sposób zasoby informacyjne opisujące będą przebieg siedmiu kampanii promocyjnych sklepu internetowego (księgarni internetowej) prowadzonych za pośrednictwem poczty elektronicznej w trzech okresach. Zasilona w ten sposób hurtownia będzie umożliwiała dostęp do swoich zasobów aplikacjom analityczno-raportującym, tj. Microsoft SQL Server Analysis Services, Visual Studio oraz Microsoft Excel, które autor wykorzystał w celu uzyskania wyników badania. Zastosowane aplikacje umożliwiają przegląd danych w perspektywach dostosowanych do potrzeb użytkowników oraz dokonywanie operacji z zakresu analizy wielowymiarowej.

Tematyczna hurtownia danych na potrzeby badania skuteczności kampanii promocyjnych w Internecie zostanie stworzona na podstawie modelu pojęciowego (konceptualnego) i logicznego. Na ich podstawie możliwe jest wygenerowanie modelu fizycznego hurtowni, który zostanie zaimplementowany w wybranym systemie baz danych [Todman 2005].

Model pojęciowy (konceptualny) – jego tworzenie polega na ogólnym opisie, charakterystyce i analizie modelowanego obszaru działalności przedsiębiorstwa pod kątem budowy struktur, zawartości oraz wykorzystania hurtowni tematycznej. Na tym etapie definiowane są cele biznesowe, które mają zostać osiągnięte dzięki analizie danych zgromadzonych w hurtowni. Docelowi użytkownicy określają więc swoje potrzeby informacyjne oraz oczekiwania w zakresie analizy danych. Model pojęciowy określa i definiuje zakres danych, które zasila hurtownię, oraz sposoby zarządzania i udostępniania danych użytkownikom. Pokazuje więc, jakie dane znajdują się w hurtowni oraz jak przełoży się to na układ poszczególnych tabel faktów i wymiarów, przy założeniu, że hurtownia będzie miała strukturę relacyjną.

Model logiczny – stanowi podstawę określenia schematu bazy danych będącej podstawą hurtowni. Zawiera szczegółową charakterystykę faktów, miar, wymiarów i atrybutów wymiarów, a także opisuje rodzaje i charakter powiązań między danymi. Na tym etapie projektowane są tabele faktów i wymiarów wchodzące w skład hurtowni oraz łączące je relacje. Ponadto określane są też nazwy poszczególnych pól w tabelach, typy znajdujących się w nich danych, definiowane są klucze itd.

2. Budowa modelu pojęciowego tematycznej hurtowni danych

2.1. Charakterystyka obszaru (procesu) biznesowego do modelowania

Rozważania na temat koncepcji modelu tematycznej hurtowni danych wspomagającej badanie skuteczności kampanii promocyjnych należy rozpocząć od scharakteryzowania rzeczywistości biznesowej, którą będzie opisywać model. Dlatego też autor stworzył studium przypadku odnoszące się do realnej sytuacji rynkowej. Podmiotem, który będzie użytkownikiem projektowanej hurtowni tematycznej, będzie firma Book-Net Sp. z o.o. Przedsiębiorstwo to prowadzi sklep internetowy, w którego ofercie znajdują się przede wszystkim książki i czasopisma. Księgarnia internetowa Book-Net jest serwisem transakcyjnym o zasięgu ogólnopolskim.

Firma Book-Net od wielu lat prowadzi różnego typu działania promocyjne za pośrednictwem Internetu. Ze względu na fakt posiadania długiej listy adresów internetowych osób zarejestrowanych w sklepie najbardziej odpowiednią formą promowania księgarni oraz jej oferty wydają się kampanie reklamowe za pośrednictwem poczty elektronicznej (kampanie e-mailingowe). Kampanie tego typu polegają na wysyłaniu reklamowych listów elektronicznych na indywidualne adresy e-mail odbiorców promocji.

Podstawową kwestią przy stosowaniu tej formy promocji jest zatem stworzenie bazy użytkowników, którzy będą odbiorcami komunikatów promocyjnych. W przypadku Księgarni Book-Net baza adresowa składa się z adresów e-mail prawie 150 tys. odbiorców, w tym ok. 100 tys. użytkowników z zarejestrowanym kontem w księgarni. Osoby te w procesie rejestracji podały nie tylko swój e-mail, ale także wybrane dane osobowe. Dlatego też każdy taki użytkownik opisany jest przez cechy o cha-

rakterze geograficznym, np. miejscowość, województwo, demograficznym, np. wiek, płeć, zawód, oraz psychograficznym, np. zainteresowania. Cechy mogą być bardzo przydatnymi informacjami dla sklepu, ponieważ na ich podstawie można dokonywać segmentacji konsumentów. W stosunku do odbiorców z różnych segmentów możliwe będzie więc dostosowywanie tematyki i treści wiadomości reklamowych w zależności od zadeklarowanych cech indywidualnych.

Wysyłane do adresatów wiadomości promocyjne złożone są z pola z tytułem wiadomości oraz pola zawierającego nazwę lub opis nadawcy. Wewnątrz wiadomości znajduje się forma promocyjna, która ma określony format, szatę graficzną, czcionkę itp. Zawiera ponadto odnośniki, tj. linki przenoszące na konkretną stronę oraz tekstowy komunikat reklamowy. Zawartość pól z tytułem i nazwą nadawcy wiadomości ma zachęcić adresatów do jej otwarcia, natomiast obiekty i hasła promocyjne mają za zadanie zmobilizować internautów do reakcji, która polega na kliknięciu na nie.

Przed każdą kampanią stawiane są określone cele, które przedsiębiorstwo ma zamiar osiągnąć. Cele mogą mieć charakter wizerunkowy, co jest związane ze zwiększeniem znajomości firmy i jej produktów wśród odbiorców promocji lub utrwalenia przekazu reklamowego w ich świadomości. W takiej sytuacji za osiągnięcie celu można uznać już samo otwarcie wiadomości i zapoznanie się z jej treścią. Należy jednak zauważyć, że przed promocją prowadzoną z użyciem poczty elektronicznej stawiane są raczej cele typu *direct response*, nazywane również sprzedażowymi. Ich osiągnięcie nie musi koniecznie polegać na zakupie produktu na stronie, ale powinno być związane z wykonaniem przez internautę jakichś działań po otwarciu e-maila, np. kliknięcia na odnośnik znajdujący się w obszarze formy promocyjnej. Spowoduje to przejście zainteresowanego użytkownika na stronę docelową, gdzie może on dokonać różnego rodzaju działań, takich jak np. wypełnienie ankiety, pobranie na stronie dodatkowych materiałów, zarejestrowanie się w serwisie poprzez założenie konta, rezerwacja lub zakupu produktu w sklepie itp.

Kampanie e-mailingowe prowadzone przez firmę Book-Net prowadzone są w miarę regularnie i związane są z takimi zdarzeniami, jak okresowe promocje na wybrane produkty, wprowadzanie nowej oferty produktowej, organizowanie konkursów, przeprowadzanie wśród użytkowników badań ankietowych dotyczących oceny księgarni i jej produktów. Kierownictwo firmy Book-Net zdaje sobie sprawę, że prowadzone do tej pory akcje promocyjne za pośrednictwem poczty elektronicznej opierały się w dużej mierze na intuicji i kreatywności pracowników działu reklamy. Kadra zarządzająca nie podejmowała decyzji dotyczących nowych akcji promocyjnych na podstawie szczegółowych raportów na temat skuteczności przeprowadzonych w przeszłości działań. Zarządzający opierali się jedynie na ogólnych spostrzeżeniach dotyczących wyników poszczególnych kampanii. Działania promocyjne poprzez pocztę elektroniczną prowadzone są przez firmę na coraz większą skalę, co wynika z bardzo dużej liczby adresatów. Podejmowanie przemyślanych i uzasad-

nionych ruchów w obszarze reklamy e-mailingowej wymaga jednak informatycznego wsparcia procesów szacowania efektów kolejnych kampanii promocyjnych. Niezbędne jest zgromadzenie gigantycznej ilości danych na temat internautów mających styczność z wiadomościami promocyjnymi, ich zachowań i aktywności. Uzyskanie tego typu zasobów informacyjnych jest możliwe, jednakże wymaga importu danych z różnych źródeł, takich jak serwery poczty elektronicznej, serwery serwisu internetowego księgarni, a także istniejące w przedsiębiorstwie bazy danych.

2.2. Użytkownicy docelowi hurtowni i uzasadnienie biznesowe jej budowy

Użytkownikami informacji, które będą pochodziły z projektowanej hurtowni danych, będą przede wszystkim osoby odpowiedzialne za planowanie, wykonanie i koordynowanie poszczególnych kampanii e-mailingowych. Fakt posiadania dostępu do dokładnych danych na temat przebiegu i skuteczności kampanii w różnych układach i na dowolnych poziomach szczegółowości daje bowiem możliwość wyciągania prawidłowych wniosków w przyszłości, a co za tym idzie, optymalizowania prac nad kolejnymi kampaniami tego typu.

Projektowana hurtownia będzie stanowiła źródło cennych informacji także dla kadry zarządzającej koordynującej działania promocyjne w Internecie z wykorzystaniem różnych instrumentów i narzędzi.

Formułując uzasadnienie biznesowe dla budowy opisywanej hurtowni, należy określić najważniejsze efekty analiz, które będą możliwe do przeprowadzenia na podstawie zgromadzonych w niej danych. W tym celu najlepiej jest sformułować przykładowe zapytania, które decydenci, menedżerowie lub analitycy będą zgłaszać do opisywanej hurtowni tematycznej. Pytania mogą mieć następującą treść:

- Które formy reklamowe umieszczone w wiadomościach przyciągały największe zainteresowanie użytkowników i spowodowały odwiedziny sklepu internetowego?
- Jak było zainteresowanie kampanią w kolejnych dniach jej trwania oraz w poszczególnych porach dnia?
- Która forma reklamowa wykorzystana w kampanii przyniosła najlepsze wyniki w grupach odbiorców o określonych cechach demograficznych?
- Jakie są główne powody niedostarczenia wiadomości do adresatów?
- Które wiadomości mają niepoprawną strukturę i treść, co powoduje wysoki wskaźnik odrzuceń przez filtry antyspamowe?

Analiza i interpretacja zgromadzonych w hurtowni zasobów informacyjnych powinny więc pozwalać na udzielanie szybkich i trafnych odpowiedzi na pytania związane ze skutecznością danej kampanii e-mailingowej oraz jej poszczególnych elementów. Definiując główny cel budowy hurtowni, można określić go jako wzrost skuteczności prowadzonych kampanii promocyjnych prowadzonych za pośrednictwem poczty elektronicznej.

2.3. Źródła danych zasilające tematyczną hurtownię danych

Hurtownia będzie zasilana z kilku źródeł danych, którymi dysponuje przedsiębiorstwo. Pierwszym z nich będzie baza zarejestrowanych użytkowników sklepu internetowego, która zawiera adresy e-mail umożliwiające kontakt. Każdy użytkownik opisany jest ponadto przez zestaw cech geograficznych, demograficznych i psychograficznych, które go charakteryzują i umożliwiają przypisanie do określonej grupy odbiorców promocji.

Z punktu widzenia szacowania skuteczności akcji promocyjnych szczególnie ważne są źródła danych, które gromadzą szczegółowe dane z okresów realizacji kolejnych kampanii opisujące zachowania poszczególnych odbiorców. Przebieg kampanii oraz ich poszczególne parametry, poczynając od momentu rozpoczęcia, aż do momentu zakończenia, rejestrowane są przez serwery poczty elektronicznej, z której wysyłane są wiadomości reklamowe. Dlatego też podstawowym źródłem, z którego pobierane będą dane do projektowanej hurtowni, są logi serwera poczty elektronicznej przedsiębiorstwa.

Kolejnym źródłem, z którego importowane będą dane do hurtowni tematycznej, będą logi serwerów, na których znajdują się strona księgarni lub strony utworzone specjalnie na potrzeby konkretnych kampanii promocyjnych. Programy śledzące umieszczone na stronie przedsiębiorstwa lub danej kampanii pozwalają na monitorowanie ruchu, którego źródłem są wiadomości promocyjne wysyłane na skrzynki e-mail. Monitoring witryny umożliwia uzyskanie wielu szczegółowo określonych w czasie (dokładna data i godzina) danych, takich jak: odsłony strony docelowej kampanii, odwiedziny użytkowników, którzy są adresatami wiadomości promocyjnych, długość poszczególnych wizyt, klikanie na elementy strony, wypełnienie formularza stworzonego na potrzeby kampanii, założenie konta w sklepie.

Ponadto serwery księgarni internetowej Book-Net gromadzą szczegółowe dane na temat konkretnych, tj. załogowanych, użytkowników. Dotyczy to takich działań, jak: złożenie zamówienia na dany towar, dokonanie rezerwacji produktu, wysłanie zapytania związanego z ofertą lub skomentowanie produktu. Tego typu informacje mogą być również przydatnym zasobem zasilającym hurtownię tematyczną.

2.4. Określenie faktów, miar i wymiarów dla modelowanego obszaru

Projekt i sposób modelowania danych w hurtowni powinien wynikać z przewidywanego zakresu informacji, jakich ma ona dostarczyć na potrzeby decyzyjne w przedsiębiorstwie. Konstruując model pojęciowy hurtowni, należy więc wyspecyfikować przede wszystkim wskaźniki biznesowe, które będą w późniejszym etapie analizowane w różnych przekrojach i perspektywach. Wskaźniki te w omawianym przypadku będą się wiązały, najogólniej mówiąc, z efektami, jakie przyniesie dana kampania e-mailingowa. Efekty te wyrażają się przede wszystkim zainteresowaniem odbiorców działaniami promocyjnymi przedsiębiorstwa. Dotyczy to różnego rodzaju ocze-

kiwanych przez firmę działań, których poszczególni adresaci dokonują podczas trwania akcji promocyjnej. Działania te polegają przede wszystkim na otwarciu wiadomości i na kontakcie z reklamą, a także na bezpośredniej reakcji na reklamę w postaci kliknięcia na umieszczone w e-mailu formy promocyjne, które przenoszą na stronę docelową kampanii. Z powyższego wynika, że efekty kampanii będą obiektem analizy, co oznacza, że w projektowanej hurtowni tematycznej należy je traktować jako zbiór tzw. faktów. Fakt opisuje pojedyncze zdarzenie i jest daną ilościową, która reprezentuje aktywność biznesową [Nycz 2009].

Fakty opisane są natomiast przez wymiary, które określają kontekst analizy. Można więc przyjąć, że pojedynczy wymiar jest określoną cechą, która charakteryzuje dany fakt i umożliwia powiązanie go z występującymi w modelu pojęciami. Dlatego też w rozpatrywanym przypadku fakty, tj. efekty kampanii promocyjnej prowadzonej z użyciem poczty elektronicznej, można rozpatrywać według siedmiu wyszczególnionych wymiarów.

1. Wymiar CZAS – szczegółowo umiejscawia w czasie poszczególne zdarzenia w ramach kampanii, które zostały zarejestrowane w trakcie jej trwania.

2. Wymiar WIADOMOŚĆ – charakteryzuje elektroniczne przesyłki promocyjne wysyłane w ramach kampanii i wyświetlające się w skrzynce odbiorczej internauty.

3. Wymiar FORMA REKLAMOWA – określa umieszczone w wiadomości, posiadające różne formaty i właściwości, obiekty reklamowe zawierające w sobie odnośniki (linki), których kliknięcie powoduje przejście na stronę docelową kampanii.

4. Wymiar REKLAMA – odnosi się do komunikatów promocyjnych umieszczanych w obszarze form reklamowych; zawierają hasła i/lub treści promocyjne mające na celu przyciągnięcie uwagi adresatów wiadomości.

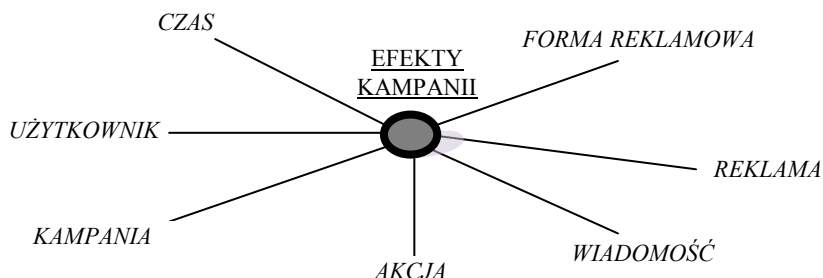
5. Wymiar KAMPANIA – określa ogólną charakterystykę kampanii, jej nazwę, cechy itp.

6. Wymiar UŻYTKOWNIK – charakteryzuje wybrane cechy adresatów kampanii promocyjnej, czyli osoby, do których wysyłane są wiadomości promocyjne.

7. Wymiar AKCJA – charakteryzuje rodzaje zdarzeń, które mogą wystąpić w okresie trwania kampanii e-mailingowej; mogą być one związane z reakcjami odbiorców (np. otwarcie, kliknięcie) lub dotyczyć akcji natury technicznej wiążących się z komunikacją pomiędzy serwerem nadawcy i skrzynką odbiorcy wiadomości (np. zwrot, odrzucenie).

Opracowywany model pojęciowy tematycznej hurtowni danych można zaprezentować z wykorzystaniem techniki modelowania punktowego (*dot modeling*) [Camilovic, Becejski-Vujaklija, Gospic 2009]. Modelowanie punktowe jest metodą pozwalającą na stworzenie ogólnego modelu pojęciowego hurtowni w postaci diagramu, który będzie obrazował sposób postrzegania obszaru do analizy z punktu widzenia wymiarów. W metodzie tej centralna część modelu, która odnosi się do zachowań, czyli faktów, jest reprezentowana przez punkt [Todman 2005]. Wymiary, według których dokonywane będą analizy umiejscowione są wokół faktu i reprezentowane są przez nazwy. Pomiędzy faktem a wymiarami znajdują się łączniki. Diagram punk-

towy projektowanej hurtowni tematycznej znajduje się na rys. 1. Przedstawia on w bardzo ogólny sposób informacje na temat pojęć, które będą podlegać późniejszej analizie.



Rys. 1. Wielowymiarowy model punktowy hurtowni tematycznej

Źródło: opracowanie własne.

Podsumowując: stworzony model pojęciowy ma strukturę gwiazdy, a wyszczególnione w nim fakty umożliwiają analizę efektów promocyjnych kampanii e-mailingowych zarówno w odniesieniu do czasu prowadzenia kampanii, używanych w nich narzędzi i komunikatów promocyjnych, jak i z punktu widzenia geograficznych, demograficznych oraz psychograficznych cech odbiorców promocji.

3. Konstrukcja modelu logicznego hurtowni tematycznej

Na podstawie modelu konceptualnego stworzony został model logiczny tematycznej hurtowni danych. Model ten będzie stanowił fundament dla schematu bazy danych, która, realizując założenia modelu pojęciowego, umożliwi jednocześnie formułowanie mechanizmów raportowania, zapytań, analiz wielowymiarowych oraz drażenia danych [Olszak 2007].

W określaniu postaci modelu logicznego zasadniczą sprawą jest zaprojektowanie struktury diagramów wchodzących w jego skład i zawierających definicje faktów, wymiarów, miar, atrybutów wymiarów i relacji pomiędzy atrybutami wymiarów.

3.1. Wybór struktury modelu logicznego

Opisywany model logiczny tematycznej hurtowni danych będzie miał wielowymiarową strukturę charakterystyczną dla schematu gwiazdy. Oznacza to, że pojedyncza tabela zawierająca dane faktograficzne otoczona będzie przez tabele referencyjne zawierające dane na temat wymiarów analizy. Tabela faktów będzie powiązana z tabelami wymiarów poprzez odpowiednie relacje pomiędzy polami. W tak określonej strukturze w następnej kolejności należy zawrzeć fakty i wymiary, które będą

istotą hurtowni tematycznej. Zarówno fakty, jak i wymiary należy umiejscowić w odpowiednio zaprojektowanych tabelach, co umożliwi wyspecyfikowanie danych niezbędnych do późniejszej analizy.

3.2. Określenie tabel wymiarów dla projektowanej hurtowni

Na etapie modelowania pojęciowego wyodrębnionych zostało siedem wymiarów, według których opisywać można przebieg i efekty poszczególnych kampanii e-mailingowych. W modelu logicznych tworzonej hurtowni dane dotyczące każdego z wymiarów zgromadzone zostaną w formie siedmiu oddzielnych tabeli wymiarów, tj.:

- 1) tabela CZAS, w której zawarte będą dane opisujące wymiar CZAS;
- 2) tabela AKCJA, w której zawarte będą dane opisujące wymiar AKCJA;
- 3) tabela UŻYTKOWNIK, w której zawarte będą dane opisujące wymiar UŻYTKOWNIK;
- 4) tabela WIADOMOSC, w której zawarte będą dane opisujące wymiar WIADOMOŚĆ;
- 5) tabela FORMA_REKLAMOWA, w której zawarte będą dane opisujące wymiar FORMA REKLAMOWA;
- 6) tabela REKLAMA, w której zawarte będą dane opisujące wymiar REKLAMA;
- 7) tabela KAMPANIA, w której zawarte będą dane opisujące wymiar KAMPANIA.

Każda tabela wymiaru zawiera kilka atrybutów opisowych, które charakteryzować będą zawarte w poszczególnych tabelach obiekty. Atrybuty są więc cechami wymiaru i zawierają dodatkowe informacje opisujące fakty. Umożliwiają ponadto logiczne pogrupowanie elementów w ramach istniejących wymiarów. Prawidłowe wyznaczenie atrybutów ma duże znaczenie dla użytkownika końcowego, ponieważ będzie on formułował zakresy interesujących go analiz według jednego lub kilku atrybutów. Odpowiednio wybrane i zdefiniowane atrybuty dają więc szansę na określenie wstępnego wyglądu raportów, sprecyzowanie perspektyw analizy oraz ustalenie jej poziomów szczegółowości. Atrybuty składają się z określonej liczby **elementów**, które zawierają w sobie wartości opisujące modelowaną rzeczywistość, takie jak poszczególne polskie wojewódzka, z których pochodzą odbiorcy kampanii promocyjnej, nazwy i rodzaje internetowych form reklamowych, miesiące i godziny prowadzenia kampanii [Januszewski 2008]. Należy zauważyć, że na etapie tworzenia raportów to właśnie według wartości atrybutów można grupować informacje w poszczególnych zestawieniach.

Poniżej przedstawiono szczegółowe specyfikacje wszystkich tabel wymiarów. Zawierają one układ pól w tabelach, informację na temat tego, czy dane pole traktowane jest jako atrybut, oraz wybrane elementy określonych atrybutów wymiarów. Ponadto w zestawieniu tabel wymiarów zaznaczone zostały (poprzez pogrubienie

Tabela 1. Tabele wymiarów nazwy pól, atrybuty oraz wybrane elementy atrybutów

Tabela	Nazwa pola	Atrybut	Wybrane elementy atrybutu
Czas	ID_CZASU	Nie	Klucz podstawowy
	ROK_NR	Tak	2012, 2013, 2014, 2015, ...
	MIESIAC_NR	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
	MIESIAC_NAZWA	Tak	styczeń, luty, marzec, ...
	TYDZIEŃ_NR	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	DZIEŃ_W_ROKU_NR	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	DZIEŃ_W_MIESIACU_NR	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	DZIEŃ_W_TYGODNIU_NR	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	DZIEŃ_NAZWA	Tak	poniedziałek, wtorek, środa, ...
	DZIEŃ_TYP	Tak	pracujący, weekend, święto
	PORA_DNIA	Tak	1 – rano, 2 – przed południe, 3 – wczesne popołudnie, 4 – wczesny wieczór, 5 – późny wieczór, 6 – noc
	GODZINA	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
	MINUTA	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	SEKUNDA	Tak	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ...
	Akcja	ID_AKCJI	Nie
KOD_AKCJI		Tak	WYS, DOST, UOTW, OTW, UKLIK, KLIK, WYP, ZALKON, ZLZAM, REZ, ZWNNS, ZWPS, ZWT, ODSPA
NAZWA_AKCJI		Tak	wysyłka, dostarczenie, unikalne otwarcie, otwarcie, kliknięcie, wypełnienie formularza, założenie konta, złożenie zamówienia, rezygnacja, zwrot niedosyt. serwera, zwrot pełna skrzynka, zwrot zerw. pol., zwrot twardy, odrzucenie spam
Uzytkownik	ID_UZYTKOWNIKA	Nie	Klucz podstawowy
	STATUS_UZYTKOWNIKA	Tak	klent, potencjalny klient, doskonał konwersji, zainteresowany ofertą, aktywny, nieaktywny
	ADRES_EMAIL	Nie	–
	SYSTEM_POCZTOWY	Tak	onet.pl, o2.pl, interia.pl, gmail.com, ...
	PLEC	Tak	K – kobieta, M – mężczyzna
	WYSZKALTALCENIE	Tak	podstawowe, zasadnicze zawodowe, średnie, politealne, ...
	ZAWOD	Tak	student/uczeń, pracownik usług i handlu, kierownik/manager, ...
	BRANZA	Tak	finanse i bankowość, edukacja/nauka, telekomunikacja/łączność, ...
	WOJEWODZTWO	Tak	dolnośląskie, śląskie, zachodniopomorskie, wielkopolskie, ...
	MIEJSCOWOSC	Tak	Wrocław, Kraków, Gdańsk, Szczecin, ...
	WIEK	Tak	18, 19, 20, 21, 22, 23, 24, 25, ...
	DATA_REJESTRACJI	Nie	–

	NAZWISKO	Nie	–
	IMIE	Nie	–
	ZAINTERESOWANIA	Tak	zdrowie, dziecko, muzyka, ...
Wiadomość	ID_WIADOMOSCI	Nie	Klucz podstawowy
	KOD_WIADOMOSCI	Tak	W_1, W_2, W_3, ...
	TYTUL_WIADOMOSCI	Tak	Zapraszamy do nas na tydzień promocyjny, Wyjątkowe oferty dla ...
	POLE_NADAWCY	Tak	Book-Net – księgarnia internetowa, Twoja księgarnia Book-Net, ...
Forma reklamowa	ID_FORMY_REKL.	Nie	Klucz podstawowy
	KOD_FORMY_REKL.	Tak	FR_1, FR_2, FR_3, ...
	KATEGORIA	Tak	tablica reklamowa, newsletter ofertowy, ...
	RODZAJ	Tak	grafika, tekst + grafika, ...
	FORMAT	Tak	HTML, JPG, ...
	CZCIONKA RODZAJ	Tak	Tahoma, Verdana, Arial, ...
	CZCIONKA ROZMIAR	Tak	12, 15, 17, 20, ...
	SZEROKOSC_OKNA	Tak	500 px, 600 px, 100%, ...
	WAGA_PLIKU	Tak	50 kB, 60 kB, 75 kB, ...
	KOLOR_TLA	Tak	granat4, zielony3, szary, ...
	GRAFIKA_W_TILE	Tak	tlo_grafika_6, promofile, ...
	OPIS	Nie	–
	Reklama	ID_REKLAMY	Nie
KOD_REKLAMY		Tak	R_1, R_2, R_3, ...
NAZWA_REKLAMY		Tak	Tydzień_Promo1, Dziecko1, ...
HASLO_REKLAMOWE		Tak	Najlepsze rady wychowawcze..., Obniżka wszystkich prod...
TEMATYKA		Tak	wszystkie kategorie, wychowanie dzieci, ...
OPIS		Nie	–
Kampania	ID_KAMPANII	Nie	Klucz podstawowy
	KOD_KAMPANII	Tak	KAMP_1, KAMP_2, ...
	NAZWA_KAMPANII	Tak	Tydzień promocyjny, Dział dziecko wersja 1, ...
	CEL1_KAMPANII	Nie	–
	CEL2_KAMPANII	Nie	–
	CEL3_KAMPANII	Nie	–
	CEL4_KAMPANII	Nie	–
	BUDZET_KAMPANII	Nie	–
DATA_ROZPOCZECIA	Nie	–	

Źródło: opracowanie własne.

i szare wypełnienie komórek) tzw. **klucze podstawowe**. W danej tabeli jest nim pole, które jednoznacznie identyfikuje wszystkie rekordy (wiersze) występujące w owej tabeli. Dane zawarte w polu oznaczonym kluczem podstawowym są więc unikalne dla danej tabeli, czyli nie ma możliwości, aby się w niej powtórzyły.

Cechą charakterystyczną każdej tablicy wymiaru jest ich względna statyczność. Wynika to z tego, że długość tabel wymiarów wzrasta niezbyt często. Tabela AKCJA jest najbardziej statyczna, ponieważ zawarta w niej lista typów zdarzeń, które mogą nastąpić podczas trwania kampanii e-mailingowej, jest kompletna z punktu widzenia potrzeb informacyjnych przedsiębiorstwa. Nie wyklucza to oczywiście zmian w przyszłości. Tablice FORMA_REKLAMOWA, REKLAMA, KAMPANIA, WIADOMOSC mogą natomiast zmieniać się o wiele częściej poprzez dopisywanie do nich dodatkowych krotek w miarę przeprowadzania kolejnych kampanii promocyjnych za pośrednictwem poczty elektronicznej. Zmiany objętości tych tabel nie będą jednak duże i polegać będą na dodaniu, w zależności od tabeli, od kilku do kilkudziesięciu krotek. Mogą być to np. nazwy nowych kampanii, nowe hasła reklamowe, tytuły wiadomości, kategorie form reklamowych itp. Znacznie szybciej może rozrastać się tablica UZYTKOWNIK, która musi być regularnie aktualizowana przez dodawanie nowo pozyskanych internautów. Najmniej statyczna tabela wymiaru to CZAS, która powiększać się będzie wraz z upływem kolejnych kampanii. Dynamika jej wzrostu zależeć będzie zatem od stopnia zainteresowania kampanią i liczbą reakcji na nią wśród jej adresatów.

3.3. Określenie struktury tabeli faktów

Konstruując tabelę faktów hurtowni tematycznej, należy przypomnieć ustalenia zawarte w modelu pojęciowym, w którym pojedynczy fakt został zdefiniowany jako efekt kampanii e-mailingowej polegający na wystąpieniu określonego zdarzenia. Tabela faktów nosić będzie nazwę „EFEKTY KAMPANII” i gromadzić będzie dane, które będą docelowo podlegać analizom za pośrednictwem narzędzi analityczno-raportujących. Zasoby informacyjne znajdujące się w omawianej tabeli będą bowiem analizowane w przestrzeni wielowymiarowej utworzonej przez poszczególne wymiary oraz znajdujące się w ich obrębie atrybuty szczegółowe.

Projektowana tablica faktów charakteryzuje się bardzo dużą objętością (liczbą wierszy) ze względu na niezwykle wysoką ziarnistość gromadzonych w niej danych. Informacje o faktach mogą być zapisywane z dokładnością do jednej sekundy. Biorąc pod uwagę, że poszczególne zdarzenia mogą dotyczyć nawet 150 tys. adresatów wiadomości promocyjnych, wielkość tabeli faktów może być ogromna. Tabela faktów „EFEKTY KAMPANII” ma zatem charakter wysoce dynamiczny, ponieważ wraz z realizowanymi przez przedsiębiorstwo kolejnymi kampaniami zapisywane są w niej nowe fakty opisujące przebieg i efekty wspomnianych działań.

Zasadniczą sprawą w projektowaniu tabeli faktów jest wyodrębnienie wchodzących w jej skład pól. W ramach tabeli EFEKTY KAMPANII ujęto następujące pola

zawierające identyfikatory atrybutów, które będą odnosiły się do tabel wymiarów. Wyróżnić można siedem takich identyfikatorów: KOD_CZASU, KOD_UZYT-KOWNIKA, KOD_WIADOMOSCI, KOD_FORMY_REKLAMOWEJ, KOD_REKLAMY, KOD_KAMPANII, KOD_AKCJI. Wszystkie pola mają postać liczbowych numerów identyfikacyjnych i odnoszą się do analogicznych pól we wszystkich tabelach wymiarów.

Zazwyczaj tabele faktów w hurtowniach tematycznych, poza identyfikatorami wymiarów, zawierają również atrybuty definiujące miary. Miary (*measures*) podobnie jak wymiary opisują fakty, a ponadto są wartościami liczbowymi przyporządkowanymi do określonego faktu. Należy też zauważyć, że wartość każdej zdefiniowanej miary uzależniona jest od zbioru wymiarów, które określają kontekst miar [Nycz 2009].

Dla przykładu dla faktu „sprzedaż” miarą może być wartość sprzedaży lub liczba sprzedanych sztuk, ponieważ określona wartościowo lub ilościowo sprzedaż jest wyznacznikiem zaistnienia faktu. Co więcej, określone w ten sposób miary zależą od wymiarów, ponieważ np. wartość sprzedaży może dotyczyć konkretnego czasu, produktu, sklepu, regionu itp. W niniejszym modelu fakty – efekty kampanii e-mailingowej – mogą być opisane przez wiele miar, które mają charakter ilościowy. Wielkości każdej z miar określają zatem w różny sposób to, jak kształtują się efekty kampanii e-mailingowej. W ramach modelu można wyszczególnić miary bezpośrednio związane z rodzajem zdarzeń, które mogą nastąpić w trakcie trwania kampanii, takich jak liczba wysłanych wiadomości, liczba dostarczonych wiadomości promocyjnych, liczba unikalnych otwarć wiadomości promocyjnych, liczba rezygnacji, liczba zwrotów itp.

W tym miejscu należy podkreślić, że rozdzielczość procesu biznesowego, jakim jest przebieg kampanii e-mailingowej, jest bardzo duża, a dane źródłowe o zainteresowaniu kampanią zapisywane są w logach serwera poczty elektronicznej z dokładnością do sekundy. Tak duża ziarnistość jest niezbędna, ponieważ dane opisujące pojedynczy fakt, czyli zdarzenie, do którego dochodzi w trakcie trwania kampanii, rejestrowane jest każdorazowo dla pojedynczego użytkownika. Każdy fakt może więc dotyczyć tylko jednego ze zdarzeń.

Na etapie określania koncepcji hurtowni tematycznej autor zauważył, że wykorzystanie w tabeli faktów miar w standardowym (powyższym) układzie wiązać się będzie ze znaczącą nadmiarowością danych w tabeli faktów. Taka sytuacja spowodowałaby, iż struktura hurtowni byłaby nieoptymalna z punktu widzenia gromadzenia danych oraz formułowania zapytań do bazy na etapie analizy. W związku z tym autor zdecydował się na wykorzystanie w projektowanej hurtowni tematycznej tabeli faktów o specyficznej strukturze, tj. tabeli typu *factless fact* [<http://www.kimball-group.com/html/designtips.html>]. Nie zawiera ona faktów określonych przez numeryczne miary, jak to jest w przypadku typowych tabel faktów. Jest ona optymalnym rozwiązaniem, idealnie pasującym do specyfiki gromadzenia danych opisujących przebieg i efekty kampanii prowadzonych za pośrednictwem poczty elektronicznej.

Tego typu tabela faktów jest przeznaczona do śledzenia różnego typu zdarzeń będących przedmiotem zainteresowania przedsiębiorstwa, które wystąpiły lub nie w danym punkcie w czasie. W projektowanej hurtowni mogą być to otwarcia wiadomości, kliknięcia na formy promocyjne, różnego rodzaju zwroty itd. Tabela ta jest więc odpowiednią formą gromadzenia informacji o faktach, które charakteryzują się maksymalną ziarnistością wystąpienia.

Mimo że w tabeli faktów nie występują określające je miary, nie wyklucza to możliwości uzyskania w procesie analizy wielowymiarowych wyników badanego zjawiska. W tabeli faktów przedstawione są bowiem relacje między atrybutami istniejących wymiarów bez wykorzystania wartości miar. Samo istnienie tego typu relacji uznawane jest za fakt. W skład tabeli faktów wchodzi zatem klucze obce wszystkich tabel wymiarów.

Tabela faktów EFEKT ma jedno specyficzne pole identyfikacyjne ID_EFEKTU, które zostało stworzone specjalnie po to, aby pełnić funkcję klucza głównego tabeli. Pole to nie odnosi się do żadnej z tabel wymiarów. W trakcie trwania kampanii promocyjnych zapisywane są w nim unikalne numery identyfikacyjne wszelkich zdarzeń związanych z wysłanymi i odbieranymi przesyłkami. Dlatego też pole to jednoznacznie identyfikuje wszystkie rekordy tabeli EFEKT. Ponadto każde pole w tablicy faktów ma typ danych reprezentujący liczbę całkowitą (int). Struktura tabeli faktów zaprezentowana została na rys. 2.

	Column Name	Data Type	Allow Nulls
▶	ID_EFEKTU	int	<input type="checkbox"/>
	ID_KAMPANII	int	<input type="checkbox"/>
	ID_CZASU	int	<input type="checkbox"/>
	ID_UZYTKOWNIKA	int	<input type="checkbox"/>
	ID_FORMY_REKLAMO...	int	<input type="checkbox"/>
	ID_REKLAMY	int	<input type="checkbox"/>
	ID_WIADOMOSCI	int	<input type="checkbox"/>
	ID_AKCJI	int	<input type="checkbox"/>
			<input type="checkbox"/>

Rys. 2. Widok projektowy tabeli faktów EFEKT

Źródło: opracowanie własne.

Wydarzenia, które znajdują się w tabeli faktów, należy traktować jako zestawienie punktów styczności kluczy wymiarów w danym punkcie czasu i przestrzeni. Tego typu podejście do tabeli faktów może być wykorzystywane do generowania różnorodnych raportów poprzez dokonywanie obliczeń dotyczących liczby wystąpień obserwowanych zjawisk za pomocą wybranych kryteriów. Jak już wspomniano, każdy fakt zapisywany w tabeli faktów może dotyczyć jednego z piętnastu zdarzeń wyszczególnionych w tab. 2.

Tabela 2. Zdarzenia możliwe do wystąpienia podczas trwania kampanii e-mailingowej

Lp.	Rodzaj zdarzenia
1	Wysłanie wiadomości na konkretny adres e-mail danego użytkownika.
2	Dostarczenie wiadomości na wspomniany adres.
3	Pierwsze (unikalne) otwarcia e-maila przez danego użytkownika.
4	Kolejne (drugie itd.) otwarcia e-maila przez danego użytkownika.
5	Pierwsze (unikalne) kliknięcia na znajdującą się w wiadomości formę reklamową przez użytkownika.
6	Kolejne (drugie itd.) kliknięcia na znajdującą się w wiadomości formę reklamową przez użytkownika.
7	Wypełnienie formularza znajdującego się na witrynie, na którą danego użytkownika przenosi kliknięta forma reklamowa.
8	Założenie przez użytkownika konta w sklepie internetowym, po wejściu z klikniętej formy promocyjnej.
9	Dokonanie w sklepie internetowym zakupu przez danego użytkownika, który wszedł na stronę sklepu za pośrednictwem formy reklamowej w wiadomości.
10	Zrezygnowanie przez danego użytkownika z subskrypcji wiadomości promocyjnych księgarni, poprzez kliknięcie odpowiedniego linka znajdującego się w wiadomości.
11	Przeniesienie wiadomości wysłanej do danego użytkownika do katalogu SPAM skrzynki e-mail.
12	Zwrot wiadomości przesłanej do danego użytkownika spowodowany niedostępnością serwera.
13	Zwrot wiadomości przesłanej do danego użytkownika spowodowany przepełnioną skrzynką odbiorczą.
14	Zwrot wiadomości spowodowany zerwaniem połączenia z serwerem odbiorcy.
15	Zwrot twardej wiadomości przesłanej do danego użytkownika.

Źródło: opracowanie własne.

	ID_AKCJI	KOD_AKCJI	NAZWA_AKCJI
▶	1	WYS	Wysyłka
	2	DOST	Dostarczenie
	3	UOTW	Unikalne otwarcie
	5	OTW	Otwarcie
	6	UKLIK	Unikalne kliknięcie
	7	KLIK	Kliknięcie
	8	WYP	Wypełnienie formularza
	9	ZALKON	Założenie konta
	10	ZLZAM	Złożenie zamówienia
	11	REZ	Rezygnacja
	12	ZWNS	Zwrot niedost. serwera
	13	ZWPS	Zwrot pełna skrzynka
	14	ZW7P	Zwrot zerw. poł.
	15	ZWI	Zwrot twardej
	16	ODSPA	Odrzucenie spam
*	WAZ	WAZ	WAZ

Rys. 3. Lista zdarzeń w formie elementów atrybutów w tabeli AKCJA

Źródło: opracowanie własne.

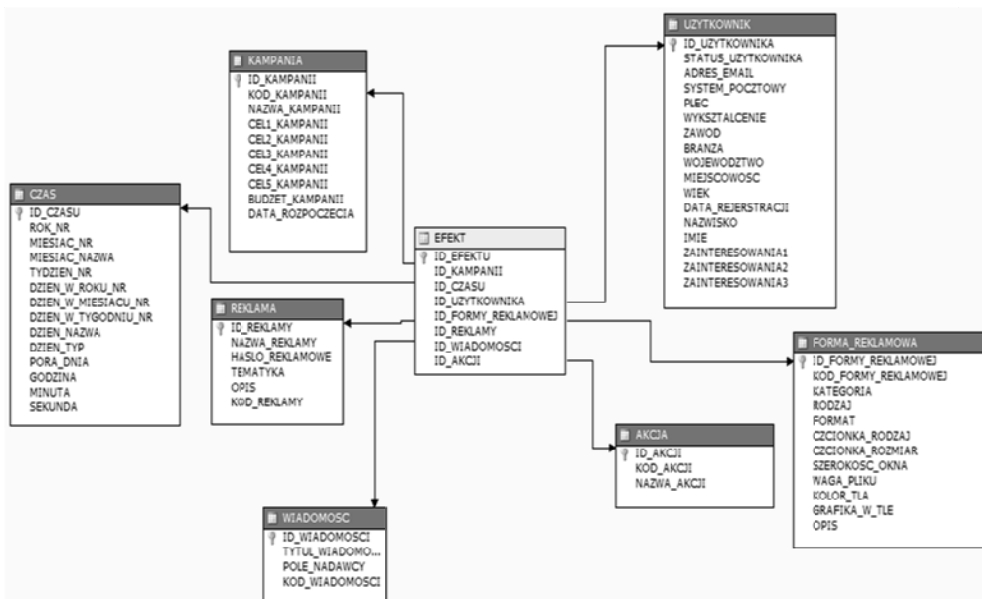
Lista możliwych do wystąpienia zdarzeń została zapisana w formie elementów atrybutów w tabeli wymiaru AKCJA, a każdemu rodzajowi zdarzenia został przypisany odpowiedni identyfikator. Pokazuje to rys. 3 przedstawiający zrzut ekranowy tabeli AKCJA.

Jedyną operacją agregującą mającą znaczenie w kontekście analizy faktów jest zatem operacja zliczania (*count*), dzięki której możliwe jest zliczanie liczby wystąpień określonych kombinacji wymiarów. Agregacja może się odbywać na dowolnych poziomach szczegółowości określonych przez hierarchię jednego lub więcej wymiarów [Januszewski 2008]. Zgromadzone w tabeli faktów dane stanowią też podstawę do obliczania wartości różnego typu wskaźników skuteczności kampanii, których wzory definiowane są na podstawie określonych metod szacowania efektów promocyjnych. Uzyskane wartości można będzie poddawać wnikliwej analizie biznesowej przez pryzmat istniejących wymiarów i atrybutów wymiarów. Wyniki obliczeń dokonywanych na faktach, zarówno wynikające z agregacji, jak i te pochodzące z podstawień do wskaźników, będą głównym rezultatem projektowanej hurtowni tematycznej.

3.4. Charakterystyka relacji łączących tabelę faktów z tabelami wymiarów

Opisując budowę tabeli faktów i wymiarów, określono, że w tablicy faktów znajdują się pola zawierające identyfikatory atrybutów wymiarów, tj. pola, które pochodzą bezpośrednio z tabel wymiarów. Oznacza to, iż pola KOD_KAMPANII, KOD_CZASU, KOD_UZYTKOWNIKA, KOD_WIADOMOSCI, KOD_FORMY_REKLAMOWEJ, KOD_REKLAMY oraz KOD_AKCJI są w tabeli EFEKT tzw. kluczami obcymi. Pola te są również kluczami podstawowymi w tabelach KAMPANIA, CZAS, UZYTKOWNIK, WIADOMOSC, FORMA_REKLAMOWA, REKLAMA, oraz AKCJA. Oparte jest to na właściwości, że klucz obcy jest polem w jednej tabeli, której wartości są określone przez klucz podstawowy z innej tabeli [Poe, Klauer, Brobst 2000]. Klucze podstawowe z tabel wymiarów łączą się więc poprzez relację z kluczami obcymi z tabeli faktów, co powoduje, że jednemu rekordowi z tabeli wymiaru może odpowiadać wiele rekordów z tabeli faktów. Relacje występujące w tematycznej hurtowni mają zatem typ jeden do wielu.

Na podstawie dokonanych założeń można skonstruować docelową strukturę modelu logicznego tematycznej hurtowni danych stworzonej na potrzeby szacowania skuteczności kampanii promocyjnych prowadzonych za pośrednictwem poczty elektronicznej. Model przedstawiony został w formie diagramu związków encji. Został on zaprezentowany na rys. 4. Model ma strukturę gwiazdzistą w pierwszej postaci normalnej. Tabela faktów połączona jest relacjami z siedmioma otaczającymi ją tabelami wymiarów. Relacje typu jeden do wielu pomiędzy tabelami zachodzą poprzez zaznaczone strzałkami pola zawierające numery identyfikacyjne atrybutów poszczególnych wymiarów.



Rys. 4. Model logiczny tematycznej hurtowni zaprojektowanej na potrzeby szacowania skuteczności kampanii promocyjnych prowadzonych za pośrednictwem poczty elektronicznej

Źródło: opracowanie własne.

Realizacja zaprezentowanego powyżej modelu logicznego w określonej bazie danych będzie tzw. **modelem fizycznym** tematycznej hurtowni danych [Surma 2009]. Model fizyczny wygenerowany na bazie modelu logicznego polega na nałożeniu logicznej struktury hurtowni na fizyczne urządzenie. Model fizyczny zawiera charakterystykę wszelkich parametrów zapewniających właściwe działanie hurtowni, tj. zasady indeksowania, kopiowania czy partycjonowania danych, jak określa szczegółową specyfikację sprzętową. Model ten zostanie zaimplementowany w konkretnej platformie bazodanowej oraz zainstalowany na serwerze. W przypadku scharakteryzowanej hurtowni będzie to Microsoft SQL Server 2008 R2.

Literatura

- Camilovic D., Becejski-Vujaklija D., Gospic N., *A call detail records data mart: Data modelling and OLAP analysis*, „Computer Science and Information Systems” 2009, vol. 6, no. 2.
<http://www.kimballgroup.com/html/designtips.html>.
- Januszewski A., *Funkcjonalność informatycznych systemów zarządzania. Systemy Business Intelligence*, Wydawnictwo Naukowe PWN, Warszawa 2008.
- Nycz M., *Techniki pozyskiwania wiedzy w hurtowni danych*, [w:] R. Knosala (red.), *Komputerowo zintegrowane zarządzanie*, t. II, Oficyna Wydawnicza PTZP, Opole 2009.

- Olszak C.M., *Tworzenie i wykorzystanie systemów Business Intelligence na potrzeby współczesnej organizacji*, Wydawnictwo Akademii Ekonomicznej, Katowice 2007.
- Poe V., Klauer P., Brobst S., *Tworzenie hurtowni danych*, Wydawnictwa Naukowo-Techniczne, Warszawa 2000.
- Surma J., *Business Intelligence. Systemy wspomaganie decyzji biznesowych*, Wydawnictwo Naukowe PWN, Warszawa 2009.
- Todman Ch., *Projektowanie hurtowni danych*, Wydawnictwa Naukowo-Techniczne, Warszawa 2005.

MODEL OF DATA MART PREPARED DUE TO RESEARCH REGARDING PROCEEDING OF E-MAIL CAMPAIGNS AND ITS EFFECTS

Summary: The article presents projects of conceptual and logistic model of data mart that can be used during research regarding/concerning the effectiveness of e-mail campaigns. Among the warehouse's main objectives there will be: gathering, unification and organization of data regarding proceeding of campaigns and its effects during various periods of time. Within conceptual model project there have been characterized the business area in preparations, target users of designed tools and data sources that contribute to the warehouse. Moreover major facts and warehouse dimensions have been assigned. During the creation of a logical model, the choice has been made regarding its structure. Facts table has been designed along with dimensions table, including its attributes and elements of these attributes. Furthermore relations between particular tables have been defined. The prototype of a data mart will be a base to conduct multidimensional analysis by using adequate tools, type OLAP. The objectives of these tools will be establishing the effectiveness of all carried out on the Internet promotional actions.

Keywords: Internet promotion, e-mail campaigns, measurement of Internet promotion effectiveness, data mart, conceptual model, logical model.