

ROBOT SIECIOWY W EKSPLORACJI INTERNETOWEGO SYSTEMU AUKCYJNEGO I OCENA HANDLU C2C W POLSCE

Artur STRZELECKI, Tomasz BACEWICZ, Marcin ŚCIAŃSKI

Streszczenie: Badanie pokazuje rezultaty ilościowej analizy rozwoju biznesu elektronicznego na przykładzie systemu aukcji elektronicznych. Dane użyte w przeprowadzonym doświadczeniu zostały pobrane z największego polskiego portalu aukcyjnego Allegro. Artykuł ukazuje dynamikę rozwoju portalu aukcyjnego oraz specyfikę zachowań oferentów i licytantów. W szczególności opracowanie skupia się na analizie zaobserwowanych trendów sprzedażowych handlu elektronicznego C2C.

Słowa kluczowe: aukcje internetowe, handel elektroniczny, C2C.

1. Wprowadzenie

Dynamiczny rozwój Internetu niesie ze sobą nowe możliwości, ale także nowe zagrożenia i wymagania. Oczywiście jest dziś, żeby przemysł rozwijał się w sposób prawidłowy, dynamiczny i konkurencyjny wymagana jest jego jak najszerza automatyzacja. Korzystając z najnowszych technologii pomiarowych automatyzowany jest niemal każdy proces produkcyjny. Wyraźnie zwiększa się jego efektywność, poprzez obniżenie kosztów pracy ludzkiej oraz wyeliminowania błędów, jakie człowiek może popełnić w procesie decyzyjnym. Podobnie należy postąpić w przypadku analizy tak ogromnego zasobu informacji, jakim jest sieć Internet i wszystkich informacji zawartych w poszczególnych witrynach internetowych. Dodanie dużych zasobów informacji do Internetu lub ich pobranie i analiza wymaga pełnej automatyzacji. Manualne zebranie i opracowanie takiego ogromu informacji jest niemożliwe do wykonania w czasie rzeczywistym. Aplikacje automatycznie przeszukujące Internet, w celu zebrania informacji wymaganych przez ich autora nazywamy Robotami Sieciowymi.

Najbardziej powszechnym rodzajem tego typu aplikacji są Roboty Indeksujące wyszukiwarek internetowych. Doskonałym przykładem może być Googlebot pracujący dla najpopularniejszej wyszukiwarki internetowej Google, który przechodzi pomiędzy poszczególnymi stronami korzystając z odnośników na nich umieszczonych. Tworzy on sieć powiązań pomiędzy stronami i na podstawie ich siły reprezentowanej wynikiem algorytmu PageRank i jego późniejszych modyfikacji przedstawia rezultaty na stronie z wynikami wyszukiwania. Googlebot został zaprogramowany w dwóch wersjach. Pierwsza z nich to deepbot, który ma za zadanie indeksować jak najwięcej informacji na użytek wyszukiwarki, drugi natomiast to freshbot, mający na celu odnalezienie w sieci nowo powstałych witryn internetowych, które przez pewien okres są w wyszukiwarce Google promowane. Robotami tego typu możemy świadomie kierować wykorzystując plik robots.txt w którym zawieramy instrukcje zezwalające na indeksowanie wybranych części witryny i blokowanie innych.

Roboty wyszukiwarek są aplikacjami użytecznymi dla całej internetowej społeczności, ułatwiają nam one szybkie odnalezienie wymaganych informacji. Niestety nie robią tego bezkosztowo, ponieważ wykorzystują pewną ilość transferu, jakim dysponuje witryna, za

który jej właściciel zmuszony jest zapłacić. Omówiony rodzaj robota jest programem pobierającym informacje z sieci, ale to tylko jeden z wielu typów robotów sieciowych.

Drugą bardzo dużą i niestety szkodliwą dla twórców stron internetowych oraz administratorów for internetowych grupą robotów są spambots. Spambots mają za zadanie przede wszystkim umieszczanie i rozpowszechnianie informacji w sieci. Dwa najpowszechniejsze typy spam botów to 'email spambot' i 'forum spambot'. Pierwszy z nich przemierza Internet w poszukiwaniu adresów e-mail, które zostają zapisane w jego bazie danych. Następnie na zebrane adresy e-mail rozestana zostaje wiadomość najczęściej zawierająca treści reklamowe lub niebezpieczne odnośniki prowadzące do zainstalowania niepożądanego oprogramowania na naszym komputerze. Problemy, jakie wiążą się z takimi programami to generowanie niepotrzebnego ruchu w Internecie, obciążającego jego infrastrukturę oraz zasypanie naszych skrzynek mailowych niechcianymi wiadomościami. Forum spambot służy natomiast do umieszczania informacji reklamowych na forach dyskusyjnych. Roboty tego typu są w stanie przejść poprzez proces rejestracji użytkownika i w losowym temacie założonym uprzednio na forum dodawać swoje wiadomości. Zawierają one najczęściej teksty reklamowe wraz z odnośnikami do promowanych stron. Masowe umieszczenie hiperłącza do wybranej strony na wielu forach prowadzi do wzrostu wartości PageRank w wyszukiwarce Google, pozwalając wypromować w ten sposób witrynę internetową, skutkując jednocześnie przekłamaniem realnej wartości strony i trafności wyszukiwania. Programami typu spambot nie możemy w żaden sposób kierować, jedyny wpływ, jaki na nie mamy to próba ich zablokowania, poprzez zablokowanie adresów e-mail, z których wysyłane są wiadomości lub z których wykorzystaniem poszczególne roboty rejestrują się jako użytkownicy forum. Dodatkowo w procesie rejestracji wymaga się odczytania grafiki zawierającej ciąg znaków zapisany w sposób utrudniający podejrzenia go za pomocą systemów rozpoznających znaki. W sieci istnieją również strony generujące masowe ilości losowych (nieistniejących) adresów e-mail, zawierających jednocześnie odnośniki do swoich kopii, skutecznie zapełniając bazę danych robotów bezużytecznymi informacjami.

Przyglądając się sposobowi przepływu danych w globalnej sieci, oraz wykorzystaniu robotów sieciowych możemy intuicyjnie dokonać pewnego porównania Internetu do wielkiego rynku informacji. Na nim pewne zbiory cennych danych są strzeżone, a dostęp do nich zostaje ograniczony lub staje się odpłatny. Inne natomiast za pomocą organizacji marketingowych jakimi niewątpliwie stają się roboty sieciowe zostają wypromowane. Za pomocą robotów sieciowych można zebrać interesujące nas informacje na temat wybranej grupy społecznej i skutecznie wykorzystać je w celu promowania lub rozwoju witryny internetowej. Należy pamiętać jednak, że ocena sposobu owej promocji oraz konkurencji zawsze styka się z tematem etyki działania w Internecie. Budując aplikacje sieciowe należy pamiętać o tym, aby ich działanie nie było uciążliwe dla współużytkowników Internetu. Nie należy naruszać ich prawa do prywatności, a jeżeli sami posiadamy ważne informacje powinniśmy zadbać o ich poprawne zabezpieczenie.

Wzrost liczby zarejestrowanych użytkowników oraz przeprowadzonych przez nich transakcji jaki obecnie notujemy pokazuje zwiększającą się popularność internetowych systemów aukcyjnych. Zaletą handlu elektronicznego C2C jest atrakcyjność cen, szeroka oferta asortymentu, możliwość szybkiego odnalezienia produktu i wygoda zawierania transakcji. Oprócz niezawodności systemu aukcyjnego od strony technicznej, fundamentalne znaczenie ma zaufanie kupującego do sprzedawcy, który swą wiarygodność potwierdza otrzymanymi komentarzami.

Każda jednostka funkcjonująca w gospodarce rynkowej i dokonująca transakcji

wymiany dąży do realizacji pewnych celów. Ze strony sprzedającego, który kieruje się kryterium rentowności, to maksymalizacja zysku. Sprzedawca kieruje swój towar do określonej grupy docelowej, wierząc w trafność podjętych do tej pory decyzji, dotyczących m.in. ceny oraz zakresu działalności. Dopiero po dokonaniu się transakcji kupna-sprzedaży rynek daje odpowiedź, na ile podjęte przez sprzedawcę działania okazują się rentowne [1]. Transakcje zawierane obecnie za pośrednictwem Internetu dają przestrzeń do rozwoju nowej gałęzi handlu, jaką jest e-retailing [2], czyli elektroniczna sprzedaż detaliczna. Trzeba pamiętać o drzemającym w tym rynku potencjale do wzrostu, z uwagi na powiększającą się liczbę użytkowników internetu.

W następnej części artykułu omawia pokrewne prace z dziedziny aukcji elektronicznych. W części trzeciej autorzy przedstawiają sposób pozyskania danych do analizy. W czwartej części prezentują otrzymane wyniki. W ostatniej części wskazano na wnioski i dalsze kierunki rozwoju badań.

2. Przegląd dziedziny

Dotychczas pojawiło się kilka prac, w których autorzy badali różne aspekty związane z działaniem systemów aukcyjnych. Zdonek i Zdonek [3] z Politechniki Śląskiej zbadali transakcje aukcyjne dotyczące głównie sektora B2C. Zwrócili uwagę na skuteczną konstrukcję komunikatu aukcyjnego, który w wyniku działania na odbiorcę nadawca zamierza zrealizować. Próbę przeprowadzili na wybranych czterech branżach aukcji internetowych i badali zaangażowanie kupującego, które przede wszystkim wynika z ponoszonego ryzyka finansowego oraz ryzyka społecznego. Wymieniony problem rozwiązywali wyłącznie z punktu widzenia nadawcy komunikatu, a więc osoby sprzedającego.

Chmielarz [4] z Uniwersytetu Warszawskiego przeprowadził analizę porównawczą systemów aukcyjnych w Polsce. Scharakteryzował obecny stan i rozwój polskojęzycznych aukcji internetowych. Badaniem objął dziewięć najpopularniejszych systemów aukcyjnych działających w Polsce w ówczesnym okresie. Posługiwał się danymi zawartymi w ich witrynach i funkcjonalnością samych witryn dla wyrażanych o nich opinii. W trakcie analizy porównawczej wziął pod uwagę następujące kryteria: zakres działania, usługi dodatkowe, wizualizacja, funkcjonalność ogółem, sposób płatności w portalu i prowizję od sprzedaży. Oceny dokonał dobrany zespół ekspertów.

Od dłuższego czasu badania nad reputacją uczestników aukcji prowadzi Morzy [5] z Politechniki Poznańskiej. W swoich pracach publikuje badania eksploracji danych na podstawie niewielkiej próbki zakończonych aukcji. W ostatniej pracy razem z Wierzbickim zaprezentował miary reputacji pozytywnej i negatywnej sprzedawców opartej o zmodyfikowany algorytm PageRank. Drugim środowiskiem prowadzącym badania uczciwości sprzedawców są Piasecki i in. [6] z AGH w Krakowie. W swojej pracy, przy udziale studentów, prognozują zachowania sprzedawców w oparciu o zmiany częstości wystawianych przedmiotów na aukcji oraz ceny tych przedmiotów. Przy pomocy metody drzew decyzyjnych są w stanie z 85% prawdopodobieństwem trafności wskazać sprzedawców, którzy popełnią oszustwo w trakcie transakcji. Morzy pracuje na zestawie danych otrzymanych bezpośrednio od największego polskiego serwisu aukcyjnego Allegro, natomiast Piasecki i in., pozyskali dane poprzez ekstrakowanie istotnych informacji prosto ze stron internetowych z aukcjami tego samego portalu aukcyjnego.

Tu [7] uznał, że w istniejącej literaturze nie można znaleźć wystarczającej ilości informacji na temat strategii sprzedażowych wykorzystywanych na portalu aukcyjnym

Ebay. Stworzył więc robota z wykorzystaniem języka skryptowego PHP, dzięki któremu miał możliwości zautomatyzowanego zbadania aukcji internetowych na których sprzedawano konsole XBOX. Porównując oferty oraz wyniki jakie udało się uzyskać sprzedającym wywnioskował, że darmowa przesyłka oraz dodatkowe wyposażenie najbardziej wpływają na skuteczne zakończenie aukcji, skutkujące zakupem z najwyższą ceną za towar.

Autorzy [8] przeprowadzili badania wyodrębniając z ogólnodostępnych danych prawdziwości, które mogłyby świadczyć o pewnych negatywnych wzorcach zachowań sprzedawców. Pozyskali wszystkie wystawione komentarze i profile użytkowników portalu Allegro, co przybliżyło ich do osiągnięcia rzetelnych rezultatów opartych na dużej próbie.

3. Metoda badawcza

Autorzy w analizie struktury sprzedażowej portalu aukcyjnego Allegro również wykorzystali robota sieciowego, który w pełni zautomatyzowany pobrał wiele cennych informacji o zakończonych aukcjach. Celem pracy jest zbadanie proporcji pomiędzy podażą i popytem w portalu oraz zachodzących tendencjach przy transakcjach. Wykorzystano informacje o poszczególnych narzędziach sprzedażowych dostarczanych przez portal, takich jak aukcje wieloprzedmiotowe, aukcje kup teraz, licytacja, cena minimalna oraz aktywności poszczególnych grup użytkowników - kupujących i sprzedających. Zmierzono obrót najaktywniejszych kupujących i sprzedających w przyjętym okresie czasu. Analizy są dzielone w poszczególnych kategoriach sprzedażowych, również z podziałem na kraje, w których aukcje zostały zawarte, ponieważ portal Allegro ma swoje zagraniczne filie w Bułgarii, Czechach, Rosji, Rumunii, Słowacji, Ukrainie i Rosji. Poddawane analizie aukcje są aukcjami zakończonymi, w których cena sprzedaży została wyraźnie wyznaczona oraz wiadome jest, kto został zwycięzcą aukcji. Podejście wykorzystywane w prowadzonych badaniach to podejście ilościowe. Zebrano i analizowano tylko wartości liczbowe bez udziału jakościowych aspektów aukcji. Aspektami jakościowymi są opis, zgodność towaru z opisem, czas oczekiwania na towar lub sposób i poprawność zawarcia transakcji. Informacje istotne z punktu widzenia autorów to cena sprzedaży, data sprzedaży, zwycięzca aukcji, ilość przedmiotów wystawionych, ilość przedmiotów sprzedanych w aukcji oraz kraj i branża, w których aukcja została zawarta.

Do budowy robota kolekcjonującego informacje o aukcjach dokonywanych za pośrednictwem portalu internetowego Allegro.pl wykorzystano język skryptowy PERL. Autorzy zdecydowali się wybrać to narzędzie z kilku powodów. Wcześniejsze doświadczenia przy budowie pajaków internetowych pokazały szybkość i stabilność działania skryptów napisanych przy użyciu języka PERL. Dodatkowo, obsługa wyrażeń regularnych, które są podstawą działania programu, została wdrożona w wydajny sposób, pozwalający wygodnie manipulować tekstem. Środowisko ECLIPSE umożliwia pełną integrację ze wspomnianym językiem, począwszy od kolorowania składni w czasie edycji kodu, aż po debugowanie aplikacji i testowanie wyrażeń regularnych.

W początkowej sekcji kodu działania robota następuje dołączenie bibliotek udostępniających wykorzystywane funkcje. Do tych bibliotek należą:

- LWP::RobotUA,
- DBI.

Biblioteka RobotUA pozwala dokładnie określić tożsamość oraz zachowanie robota w sieci. Wywołując metodę new() przekazujemy dwa parametry: nazwę i rodzaj agenta,

który łączy się z serwerem oraz adres internetowy, prowadzący do witryny, która może informować na przykład o pochodzeniu robota i jego przeznaczeniu.

Wykorzystując metody `delay()` oraz `timeout()` autorzy określili kolejno częstotliwość, z jaką robot będzie wysyłał żądania na zdalny serwer, a także czas, po którym żądanie będzie wygasło.

Samo żądanie wysyłane było za pośrednictwem metody `get()`, która w efekcie swojego działania umieszcza informacje o rezultacie operacji w polu `is_success`. W przypadku pomyślnego zakończenia operacji, odpowiedź serwera znajduje się w polu `content`. W omawianym przypadku podanie adresu URL jako argumentu metody `get()`, skutkowało zwroceniem przez serwer kodu źródłowego strony internetowej.

Kluczową kwestią było zapewnienie połączenia aplikacji z serwerem bazy danych. Do zrealizowania skutecznego połączenia użyto biblioteki DBI, z uwagi na to, iż udostępnia ona interfejs do wygodnej obsługi bazy danych, niezależnie od jej rodzaju. Wspomniane metody jedynie pośredniczą pomiędzy skryptem a sterownikami bazy danych. Najważniejszą z nich jest metoda `connect()` przyjmująca jako parametry kolejno:

- ciąg znaków o postaci `dbi:nazwaSterownika:nazwaBazyDanych:nazwaHosta` opisujący źródło danych,
- nazwę użytkownika bazy danych,
- hasło.

Opisywana metoda w przypadku pomyślnego połączenia zwraca uchwyt do źródła danych, który udostępnia metody pozwalające wykonywać zapytania bazodanowe. Autorzy wykorzystali metody takie jak `prepare()` oraz `execute()`. Pierwsza z nich pozwala „przygotować” zapytanie. Polecenie `INSERT` realizowane za pośrednictwem biblioteki DBI wymaga podania jedynie liczby wartości, które wstawia się do bazy. Przykładowe polecenie wygląda następująco:

```
$sth = $dbh->prepare("INSERT INTO nazwaTabeli VALUES (?, ?, ?, ?, ?) ")
```

gdzie: zmienna `$sth` przechowuje uchwyt do przygotowanego zapytania, `$dbh` to nazwa zmiennej przechowującej uchwyt do źródła danych, a znaki zapytania to informacja o liczbie wartości. Konkretnie wartości przekazujemy jako parametry metody `execute()`, której wywołanie, dla zgodności z poprzednim przykładem, wyglądałoby następująco:

```
$sth->execute( $id, $userId, $price, $quantity, $offerDate, $bought )
```

Istotnym jest, aby liczba argumentów metody `execute()` była zgodna z liczbą znaków zapytania przekazanych metodzie `prepare()`. Do przechowywania zebranych informacji o aukcjach autorzy zdecydowali się wykorzystać darmowe rozwiązanie w postaci bazy danych MySQL. Baza składa się z dwóch tabel.

Pierwsza z nich, tabela aukcje, zawierała następujące pola:

- identyfikator aukcji,
- identyfikator sprzedawcy,
- numer kategorii, do której należy licytowany przedmiot,
- typ aukcji, gdzie cyfra 0 oznacza aukcję typu „kup teraz” a cyfra 1 – licytację,
- liczbę sztuk oferowaną podczas aukcji,
- liczbę sprzedanych przedmiotów,
- liczbę pozostałych przedmiotów,
- liczbę złożonych ofert w aukcjach typu licytacja,
- oznaczenie rodzaju aukcji, gdzie cyfra 0 oznacza aukcję jednop przedmiotową a cyfra 1 – aukcję wielop przedmiotową,
- najwyższą zaoferowaną cenę,

- cenę przedmiotu, z jaką zakończyła się aukcja,
- informację, czy została osiągnięta cena minimalna, jeśli taka była ustawiona, gdzie cyfra 0 oznacza zakończenie aukcji przed osiągnięciem ceny minimalnej, a 1 - pomyślne osiągnięcie ceny minimalnej,
- datę zakończenia aukcji,
- kraj, z którego pochodzi sprzedawca, gdzie cyfry kolejno oznaczają:
 1. Polska
 2. Bułgaria
 3. Czechy
 4. Rosja
 5. Rumunia
 6. Słowacja
 7. Ukraina
 8. Węgry

Tabela oferty przechowywała z kolei:

- identyfikator aukcji,
- identyfikator kupującego,
- oferowaną kwotę,
- liczbę nabywanych sztuk,
- datę złożenia oferty,
- informację o ostatecznym zakupieniu przedmiotu, gdzie cyfra 0 oznaczała ofertę, która nie zakończyła się nabyciem przedmiotu przez danego użytkownika a cyfra 1 - oznaczała ofertę, która doprowadziła do nabycia przedmiotu.

Kod programu robota sieciowego dzieli się na trzy sekcje. W pierwszej z nich następuje dołączenie wykorzystywanych bibliotek oraz zdefiniowanie zachowania robota i nawiązanie połączenia z bazą danych, co zostało już opisane. Druga część zawiera definicje funkcji pomocniczych, wykorzystywanych w głównej pętli skryptu, dla zwiększenia czytelności kodu. Funkcje te to kolejno:

`getContent()` - funkcja wysyłająca na serwer żądanie pobrania kodu źródłowego strony.

Jako parametr przyjmuje numer aukcji, który zostaje wykorzystany do zbudowania adresu URL pobieranej strony. Zwraca odpowiednią wartość w zależności od tego, czy pobranie kodu źródłowego powiodło się, czy znaleziono aukcje o podanym numerze lub czy aukcja została dodana do archiwum aukcji portalu Allegro.

`getOffers()` - funkcja przyjmująca jako argumenty kod źródłowy pobranej strony, a także identyfikator aukcji. Funkcja na początku sprawdza, czy użytkownicy złożyli jakiegokolwiek oferty do danej aukcji, w przeciwnym wypadku kończy swoje działanie. Następnie poszukuje ofert wycofanych. W czasie przeglądania aukcji okazało się bowiem, że istnieją oferty, które zostały usunięte z licytacji na prośbę oferenta. W dalszej kolejności funkcja pobiera informacje o ofertach, które zakończyły się kupnem przedmiotu. Ostatecznie, po zapisaniu pobranych informacji do bazy danych, zwraca tablicę z kwotami poszczególnych ofert.

`convertDate()` - funkcja zamieniająca trzyliterowe oznaczenie nazwy miesiąca na tradycyjne oznaczenie, gdzie styczeń to numer 01, a grudzień to numer 12. Funkcja wykorzystuje prostą konstrukcję `switch-case` i zwraca pożądaną wartość.

`chooseCountry()` - funkcja rozpoznająca kraj, z którego pochodzi sprzedawca. Na stronie aukcji, jeśli ta jest wystawiona przez sprzedawcę pochodzącego z kraju innego niż Polska, widnieje ilustracja odpowiedniej flagi. Po sprawdzeniu nazwy ilustracji w kodzie

źródłowym strony, funkcja zwraca odpowiednie oznaczenie kraju.

Trzecia sekcja skryptu rozpoczyna się przypisaniem do zmiennej numeru aukcji, od której rozpoczyna się proces pobierania danych. W dalszej kolejności znajduje się główna pętla programu, która pobiera zawartość kolejnych stron zawierających aukcje, aż do momentu napotkania strony portalu Allegro.pl z wygenerowanym komunikatem „Nie znaleziono aukcji”.

W omawianej pętli, w pierwszej kolejności następuje sprawdzenie, czy pobranie zawartości strony powiodło się. W przeciwnym wypadku następuje powrót na początek pętli i próba ponownego pobrania kodu źródłowego. Takie rozwiązanie gwarantuje ciągłość działania skryptu w wypadku krótkotrwałych problemów z łącznością z serwerami portalu. Przejście na początek pętli, tym razem jednak wraz ze zwiększeniem licznika i pobraniem kolejnej strony, następuje po napotkaniu komunikatu, iż dana aukcja została przeniesiona do archiwum. W takim wypadku niemożliwe jest pobranie informacji na temat jej przebiegu i stanu, w jakim została zakończona.

W trakcie prac okazało się, że niektóre spośród zakończonych aukcji przeprowadzone zostały na portalach innych niż Allegro.pl, ale należących do grupy Allegro. Razem z portalem współpracują witryny otomoto.pl, otodom.pl oraz pracaallegro.pl. Autorzy analizowali aukcje przeprowadzone wyłącznie na portalu allegro.pl, dlatego wymienione wyżej przypadki zostały pominięte.

Skrypt został skonstruowany w taki sposób, aby pomijać także witryny informujące o aukcji usuniętej przez pracownika allegro.pl, bądź aukcje dostępne wyłącznie osobom, które ukończyły osiemnasty rok życia. Po wykluczeniu powyższych przypadków pozostawało sprawdzenie, czy aukcja została już zakończona oraz czy nie została przeprowadzana na serwisie aukcyjnym molotok.ru – rosyjskim odpowiedniku allegro.pl.

Wylimitowanie niepożądanych przypadków prowadzi do punktu, w którym przy pomocy metody chooseCountry() sprawdzany jest kraj, z którego pochodzi sprzedający, a także pobierana jest aktualna cena przedmiotu. Zarówno cena przedmiotu, jak i wysokość poszczególnych ofert mogła być podana w walucie obcej. Robot rozpoznaje podobne sytuacje i pobiera wyłącznie kwotę przeliczoną na polskie złote, a umieszczoną w nawiasie pod ceną pierwotną, podaną na przykład w czeskich koronach.

Następnie funkcja convertDate() pobiera informację o dacie zakończenia aukcji. Dalej kolejno uzyskuje numer kategorii, do której należy licytowany przedmiot, identyfikator sprzedawcy oraz typ aukcji (licytacja, bądź aukcja „kup teraz”). Po sprawdzeniu, czy osiągnięta została cena minimalna, jeśli taka była ustawiona, skrypt rozpoznaje rodzaj aukcji, to znaczy, czy jest to aukcja jedno lub wielopremiotowa. W tym miejscu następuje także wykorzystanie funkcji getOffers() i pobranie wszystkich złożonych ofert, jeśli tylko takie występują. Po określeniu liczby licytowanych przedmiotów i obliczeniu liczby przedmiotów sprzedanych, następuje zapisanie informacji do odpowiedniej tabeli w bazie danych, zwiększenie licznika pętli (będącego jednocześnie numerem aukcji) i przejście do kolejnej iteracji.

4. Wyniki badań

W trakcie działania robota sieciowego przez okres 14 dni zebrano następujące informacje o aukcjach.

W tabeli 1 przedstawiono ogólną charakterystykę zebranej kolekcji, w sekcji 3 przedstawiono szczegółowo, które informacje o zakończonych aukcjach internetowych znalazły się w bazie. Zebrane dane pozwoliły na przedstawienie następujących wyników badań.

Tab. 1. Ogólna statystyka zebranej kolekcji

Parametr	Wartość
Liczba zakończonych aukcji	9 457 269
Liczba złożonych ofert	5 229 711
Liczba aukcji w których złożono oferty	1 936 261
Liczba aukcji licytowanych	4 202 710
Liczba aukcji kup teraz	5 254 559
Liczba aukcji wieloprzedmiotowych	2 312 181
Liczba aukcji jednoprzecmiotowych	7 145 088

W całej próbie tylko w 20,47% aukcji zostały złożone oferty kupna to znaczy, że co piąta wystawiona aukcja doczekała się zainteresowanie kupującego i finalizacji transakcji. Średnia liczba ofert w aukcjach, w których je złożono wynosi 2,7 złożonych ofert w jednej aukcji. W badanej próbie znalazły się aukcje prowadzone we wszystkich 8 państwach gdzie Allegro posiada swoje filie. W każdym z państw występują odrębne kategorie branżowe. W tabeli 2 przedstawiono popularność aukcji internetowych ze względu na branżę. Najpopularniejszymi branżami są Motoryzacja, Odzież oraz Książki. Najmniej popularną jest branża Nieruchomości.

Tab. 2. Rozkład aukcji w kategoriach branżowych w Polsce

Branża	Wolumen
Motoryzacja	1306040
Odzież, Obuwie, Dodatki	1140007
Książki i Komiksy	650079
Dla Dzieci	528849
Dom i Ogród	455469
Kolekcje	282840
Zdrowie i Uroda	243871
Sport i Turystyka	240526
Biżuteria i Zegarki	220923
RTV i AGD	212673
Komputery	198581
Muzyka i Instrumenty	189558
Telefony i Akcesoria	157152
Antyki i Sztuka	155434
Firma i Przemysł	141434
Filmy	82504
Gry	78898
Fotografia	69015
Pozostałe	11513
Nieruchomości	7185

W zebranej kolekcji 55,56% aukcji było typu Kup Teraz. Sugeruje to większe wykorzystanie platformy aukcyjnej do sprzedaży przez firmy, sklepy i sprzedawców hurtowych, którzy są nastawieni na jak największą i najszybszą sprzedaż. Pozostała część

44,44% to aukcje typu licytacja czyli aukcja klasyczna (angielska). Co czwarta wystawiona aukcja jest wieloprzedmiotowa, natomiast trzy na cztery aukcje oferują tylko jeden towar lub usługę dla licytanta. Sprzedawcy na aukcjach wieloprzedmiotowych oferują średnio 29 sztuk towaru lub usługi, przy czym maksymalna liczba wystawionych sztuk na sprzedaż wynosi 100. Mimo tak dużej liczby sztuk wystawianych przedmiotów i usług na sprzedaż, to średnio tylko 1,8 sztuki zostaje sprzedanych na aukcji. W zebranej próbie 408221 sprzedawców oferowało swoje towary i usługi na aukcjach, ale tylko 199650 sprzedawcom udało się sprzedać towar. Najaktywniejszy oferent „tani-antykwarjat”, w ciągu 40 dni oferował swoje produkty na 21871 aukcjach, jednak tylko 517 zakończyło się złożeniem oferty przez licytującego, co daje skuteczność sprzedaży na poziomie 2,36% sfinalizowanych aukcji. Łączna wartość sprzedanych przedmiotów przez „tani-antykwarjat” wyniosła 7 067,29 PLN, przy czym wszystkie oferowane przedmioty miały wartość 434 032,60 PLN. Najaktywniejszy sprzedawca „CostumeJewellery” w ciągu 26 dni oferował swoje produkty na 5644 aukcjach, z czego 4693 zakończyło się złożeniem oferty przez licytującego, co daje skuteczność sprzedaży na poziomie 83,15% sfinalizowanych transakcji. Łączna wartość sprzedanych przedmiotów przez „CostumeJewellery” wyniosła 98 057,52 PLN, przy czym wszystkie oferowane przedmioty miały wartość 386 899,16 PLN.

W zebranej próbie 1448401 kupujących licytowało towary i usługi na aukcjach ale tylko 1272283 z nich faktycznie wygrało aukcję. Najaktywniejszy licytujący „Truelia”, w ciągu 36 dni złożył oferty na 2447 aukcjach, jednak tylko 277 zakończyło się wygraniami aukcji przez licytującego, co daje skuteczność wygrywania aukcji na poziomie 11,32% aukcji, w których brał udział. „Truelia” wszystkie aukcje w badanym okresie wygrał na kwotę 429,53 PLN, przy czym łączna wartość złożonych ofert razem z ofertami nieskutkującymi wygraniami aukcji wyniosła 15 162,53 PLN.

Najaktywniejszy kupujący „Zoraya” w ciągu 40 dni złożył oferty na 865 aukcjach, z czego 798 zakończyło się wygraniami aukcji przez licytującego, co daje skuteczność wygrywania aukcji na poziomie 92,25% aukcji, w których brał udział. „Zoraya” wszystkie aukcje w badanym okresie wygrał na kwotę 1 539,27 PLN, przy czym łączna wartość złożonych ofert razem z ofertami nieskutkującymi wygraniami aukcji wyniosła 1 937,37 PLN.

Wśród czterech identyfikowanych typów aukcji: wieloprzedmiotowe, jednoprzedsmiotowe, kup teraz oraz licytacja, najskuteczniejszą formą sprzedaży jest aukcja wieloprzedmiotowa. Osiągnęła ona w badanej próbie skuteczność 35% złożonych ofert kupna w zakończonych aukcjach. Tylko 14,89% aukcji jednoprzedsmiotowych zakończyło się finalizacją transakcji. Aukcje typu kup teraz i licytacja osiągnęły bardzo podobną skuteczność, odpowiednio 19,96% i 19,61% zrealizowanych transakcji.

Wśród czterech połączonych typów aukcji: wieloprzedmiotowe-licytacja, wieloprzedmiotowe-kup teraz, jednoprzedsmiotowe-licytacja oraz jednoprzedsmiotowe-kup teraz najskuteczniejszą formą sprzedaży jest aukcja wieloprzedmiotowa-kup teraz. Osiągnęła ona w badanej próbie skuteczność 35,19% złożonych ofert kupna w zakończonych aukcjach. Tylko 8,42 aukcji jednoprzedsmiotowych-kup teraz zakończyło się finalizacją transakcji. Aukcje typu wieloprzedmiotowa-licytacja oraz jednoprzedsmiotowa-licytacja osiągnęły odpowiednio 25,72% i 19,54%, ale tych drugich aukcji było ponad 100-krotnie więcej od tych pierwszych. W aukcjach typu licytacja średnia liczba składanych ofert wynosi 2,95 oferty, przy czym 4,72% licytacji nie osiągnęło ceny minimalnej. Średnia wartość najwyższej złożonej oferty dla aukcji, które osiągnęły cenę minimalną wynosi 76,55 PLN, a dla aukcji, które nie osiągnęły ceny minimalnej wynosi 1353,26 PLN.

5. Podsumowanie

Analiza poszczególnych parametrów aukcji pozwala ukazać, w jaki sposób zmienia się wolumen transakcji oraz sposób sprzedaży. Istotne jest, że cały proces analizy i wnioskowania został przeprowadzony w oparciu o dane ogromnej liczby blisko 10 milionów transakcji. Jesteśmy przekonani, że dogłębna analiza posiadanych informacji rozszerza naszą wiedzę o polskim handlu w portalu aukcyjnym Allegro, dzięki której możemy z większym rozeznaniem podejść do internetowych transakcji aukcyjnych.

Literatura

1. Milewski R.: Podstawy Ekonomii. PWN, Warszawa 1999, s. 57.
2. Banasikowska J., Woźniak I.: Transakcje elektroniczne i czynniki wpływające na ich rozwój. Systemy Wspomagania Organizacji, pod red. H. Sroki, T. Porębskiej-Miącz, Akademia Ekonomiczna, Katowice, 2003, ss. 391-398.
3. Zdonek I., Zdonek D.: Determinanty skutecznej aukcji internetowej produktów o dużym zaangażowaniu kupującego. [w:] Zastosowanie systemów e-biznesu w gospodarce, pod red. W. Chmielarza, T. Parysa, Wyższa Szkoła Ekonomiczno-Informatyczna, Warszawa, 2007, ss. 123-135.
4. Chmielarz W.: Próba analizy porównawczej serwisów aukcji internetowych w Polsce. [w:] Współczesne trendy w informatyce ekonomicznej, Roczniki Kolegium Analiz Ekonomicznych 2006, nr 16, SGH, Warszawa, 2006, ss. 13-26.
5. Morzy M., Wierzbicki A.: Eksploracja aukcji internetowych w poszukiwaniu pozytywnej i negatywnej reputacji sprzedawców. [w:] Technologie Przetwarzania Danych, II krajowa konferencja naukowa pod red. T. Morzego, M. Gorawskiego, R. Wrembla, Wydawnictwo Politechniki Poznańskiej, Poznań. 2007 ss. 103-114.
6. Piasecki J., Roczniak W., Zygmunt A., Koźlak J.: Analiza wzorców nielegalnych zachowań w aukcjach internetowych. [w:] Technologie Przetwarzania Danych, II krajowa konferencja naukowa, pod red. T. Morzego, M. Gorawskiego, R. Wrembla, Wydawnictwo Politechniki Poznańskiej, Poznań, 2007 ss. 231-242.
7. Tu Y.: An application of web-based data mining: selling strategies for online auctions. Online Information Review; Volume: 32 Issue: 2; 2008 Technical paper.
8. Strzelecki A., Bacewicz T., Ściański M.: Ocena zachowania użytkowników platformy handlu C2C na podstawie eksploracji danych i ich aktywności w internetowym systemie aukcyjnym. „E-mentor” 2008, numer 2(24).

Mgr Artur STRZELECKI
Katedra Informatyki
Tomasz BACEWICZ
Marcin ŚCIAŃSKI
Koło Naukowe Informatyki „Informaticus”,
Akademia Ekonomiczna w Katowicach,
40-226 Katowice, ul. Bogucicka 3,
e-mail: strzelecki@ae.katowice.pl
tomek.bacewicz@gmail.com
marcin.scianski@gmail.com