Katedra Biosensorów i Przetwarzania Sygnałów Biomedycznych

Wydział Inżynierii Biomedycznej POLITECHNIKA ŚLĄSKA



Cyberbezpieczeństwo w medycznych bazach i systemach danych

PROJEKT

Dominika Gabor Agnieszka Radziun Kamil Suchanek

Kierunek studiów: Inżynieria Biomedyczna

Specialność: Przetwarzanie i Analiza Informacji Biomedycznej

Prowadzący projekt: Dr inż. Rafał Doniec

Spis treści:

1.	Medyczne bazy danych	2
2.	Zabezpieczenia baz danych	4
2.1.	Ograniczanie dostępu	4
2.2.	Rozwiązania kryptografii	4
2.3.	Zabezpieczenia z poziomu wiersza	5
3.	Projekt bazy danych	6
3.1.	Schemat projektu	6
4.	Implementacja projektu	7
4.1.	Wdrożenie rozwiązania	9
5.	Zabezpieczenia bazy danych	10
5.1.	Szyfrowanie danych	10
5.2.	Szyfrogramy numeru pesel	11
5.3.	Szyfrowanie dwukierunkowe	13
5.4.	Szyfrowanie symetryczne	18
6.	Podsumowanie	20

1. Medyczne bazy danych

W marcu 2009 roku Ministerstwo Zdrowia zaprezentowało "Strategię E-Zdrowie Polska na lata 2009-2015", która mówi o wykorzystaniu wielu możliwości mediów cyfrowych umożliwiających wprowadzanie, przechowywanie oraz zarządzanie danymi medycznymi.[2,3] Głównym elementem elektronicznego zapisu jest baza danych, która zapewnia:

- usprawnienie procesu wymiany informacji w ramach jednostki oraz zespołu jednostek;
- wzrost bezpieczeństwa dokumentacji medycznej;
- polepszenie identyfikacji i czytelności osób, które dokonały w niej wpisu;
- ❖ podniesienie mobilności danych (również na terenie Unii Europejskiej);
- ❖ zwiększenie ilości informacji pozwalających na postawienie diagnozy;
- natychmiastowy wgląd do historii pacjenta: poddane procedury, leki, historia choroby;
- obniżenie kosztów archiwizacji danych.[3]

Bazy danych to środowisko oprogramowania, fizyczne urządzenia oraz bezpieczeństwo i administracja umożliwiająca: wprowadzanie, przechowywanie, modyfikowanie, usuwanie oraz pobieranie w sposób uporządkowany dużych ilości informacji. Bazy danych są stworzone z tabel, które zawierają kolumny (pola) i wiersze (rekordy). Kolumny zawierają pewien szczególny rodzaj informacji (określony poprzez właściwy typ danych np. int dla liczb całkowitych), który opisany jest poprzez wiersze.[2]

Występują trzy modele bazy danych: jednorodny, obiektowy i relacyjny. W przypadku modelu jednorodnego wszystkie dane są zawarte w jednej tabeli i najczęściej wykorzystywany jest w arkuszach kalkulacyjnych np. Excel. Natomiast w modelu obiektowym informacje są przechowywane w postaci całych obiektów, z którymi związany jest ich stan i zachowanie. Model ten jest dobrze dopasowany do obiektowych języków programowania takich jak C# czy Java. Najczęściej jednak jest spotykany

relacyjny model bazy danych.[1] Umożliwia on stosowanie powiązań (relacji) pomiędzy różnymi tabelami, wśród których wyróżniamy relacje:

- jeden do jednego typ relacji, w której występuje tylko jedna tabela, gdzie jednej wartości w kolumnie, odpowiada druga wartość w drugiej kolumnie;
- ❖ jeden do wielu typ relacji, w którym jednej wartości w jednej kolumnie odpowiada kilka wartości w drugiej kolumnie;
- ❖ wiele do jednego jest to odwrotność relacji jeden do wielu;
- wiele do wielu wielu wartościom w jednej kolumnie, odpowiada wiele wartości w drugiej kolumnie.

Aby relacje między tabelami były jednoznaczne, połączone tabele muszą posiadać klucz podstawowy, który jest identyfikatorem poszczególnych wierszy. W przypadku gdy w tabeli występują wartości z drugiej tabeli (identyfikatory rekordu) wtedy taką kolumnę nazywamy kluczem obcym.[1]

Najbardziej znanymi serwerami relacyjnych baz danych są: *MariaDB*, *MySQL*, *SQL Server*, *Oracle DataBase*. Każdy z nich ma swoje unikatowe cechy lecz łączy je fakt, że wszystkie wykorzystują język SQL (Structured Query Language) czyli strukturalny język zapytań. Jego działanie polega na na wysyłaniu zapytań (poleceń, kwerend) do bazy i dostarczaniu użytkownikowi wymaganych danych jak i tworzeniu oraz przekształcaniu elementów bazodanowych takich jak tabele czy procedury.[1]

Procedury są wykorzystywane aby uniknąć pisania niezmiennie tego samego kodu poprzez zapis pewnego zapytania, które zostanie wykorzystane wielokrotnie. Możliwe jest również przekazywanie parametrów do procedury, dzięki czemu jej wynik będzie zależny od danych wejściowych, a kod stanie się bardziej uniwersalny.[2]

2. Zabezpieczenia baz danych

Bazy i systemy danych ze względu na prywatność i bezpieczeństwo informacji wymagają pewnych zabezpieczeń. Formy i realizacja owych zależą od stopnia i rodzaju szacowanego zagrożenia. Dane nie mające wtórnej wartości oraz te publiczne będą wymagały wiarygodnego i stabilnego systemu bez nacisku na zabezpieczenia, natomiast systemy operujące na danych wrażliwych mogą być celem kradzieży i manipulacji.

2.1. Ograniczanie dostępu

Silnik bazodanowy zazwyczaj spełnia funkcje typowe dla serwera i nie tylko, należy go odpowiednio skonfigurować aby nie zezwalał na każdy rodzaj połączeń. W przypadku bazy danych PostgreSQL odpowiada za to plik konfiguracyjny *pg hba.conf*.

HBA oznacza uwierzytelnienie oparte na hoście. Zawiera on rekordy ułożone kolejno wiersz po wierszu. Każdy rekord określa typ połączenia, zakres adresów IP klienta, nazwę bazy danych, nazwę użytkownika i metodę uwierzytelnienia, która ma być używana dla połączeń spełniających te parametry $_{[4]}$.

Dostęp do serwera baz danych ograniczany jest również dla przypadkowych i nieproszonych nasłuchujących oraz osób, które zdołały uzyskać dostęp bezpośredni mimo postawionych blokad poprzez szyfrowanie treści osadzonej w bazie danych oraz wymienianej z klientem w danej chwili.

2.2. Rozwiązania kryptografii

Niektóre systemy bazodanowe wspierają natywnie obsługę połączeń SSL do szyfrowania komunikacji klient-serwer. SSL jest standardową technologią zabezpieczeń służącą do ustanawiania szyfrowanego łącza między serwerem internetowym a przeglądarką. To łącze zapewnia, że wszystkie dane przekazywane między serwerem internetowym a przeglądarkami pozostają prywatne i zintegrowane.

Inną opcją jest szyfrowane danych w spoczynku. Operacja ta polega na szyfrowaniu danych w całej bazie danych, w poszczególnych tabelach albo kolumnach. Wiele systemów bazodanowych posiada pewne rozszerzenia i wbudowane funkcje kryptografii.

Ze względu na sposób realizacji i zastosowanie szyfrowanie takie można podzielić na:

jednokierunkowe:

W tym przypadku nie zamierzamy odszyfrowywać danych do wglądu. Jedynym celem jest upewnienie się, że osoba, która chce uzyskać dostęp do pewnych obszarów, zna treść zaszyfrowaną, np.: hasło.

dwukierunkowe:

Symetryczne i asymetryczne, pierwsze polega na szyfrowaniu i odszyfrowywaniu danych przy pomocy jednego klucza, natomiast drugie polega na szyfrowaniu z pomocą klucza nazywanego publicznym oraz odszyfrowywaniu z pomocą klucza zwanego prywatnym.

2.3. Zabezpieczenia z poziomu wiersza

Oprócz bardziej oczywistych metod ochrony danych, systemy bazodanowe oferują również rozwiązania na poziomie mechanizmów działania silnika. Zabezpieczenia tego rodzaju komplikują zarządzanie bazą danych.

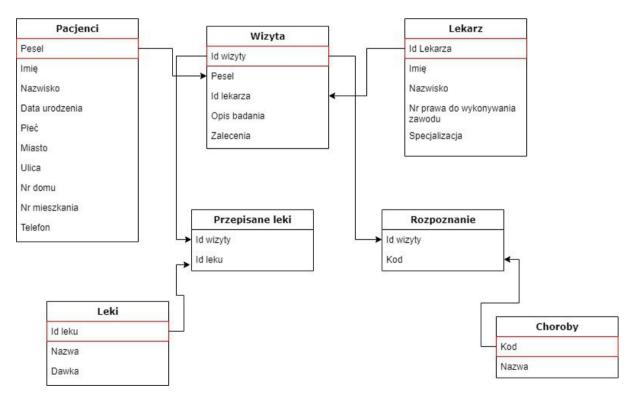
Ogólnie polega to na przyznawaniu uprawnień (GRANT) dla poszczególnych kolumn i tabel do wyświetlania rekordów, ich modyfikacji. Przykładowo standardowy klient poprzez aplikację zewnętrzną może być przyporządkowany do pewnego spektrum ról w bazie danych, mających bardzo ograniczone możliwości, natomiast inny rodzaj użytkownika z poziomu bazy danych może mieć większe uprawnienia, jednak nie stosuje komunikacji przez wspomnianą aplikację, dzięki czemu można ograniczyć możliwy, szkodliwy wpływ nieprzyjaznych zapytań do bazy danych z poziomu aplikacji klienckiej.

3. Projekt bazy danych

Zaprojektowano ideową bazę danych dla przychodni, a konkretnie części zarządzającej danymi pacjentów, związanych z wizytą. Wykorzystano tu relacyjny model bazy danych, który umożliwia przedstawienie powiązań (relacji) pomiędzy różnymi tabelami,. Relacje występujące w projekcie to jeden do wielu oraz wiele do wielu.

3.1. Schemat projektu

Projekt bazy składa się z siedmiu tablic: *Pacjenci, Wizyta, Lekarz, Leki, Choroby, Przepisane lek*i i *Rozpoznanie*. Na schemacie zaznaczono strzałkami relacje pomiędzy poszczególnymi tabelami oraz przy pomocy czerwonych ramek oznaczono klucze główne tablic.

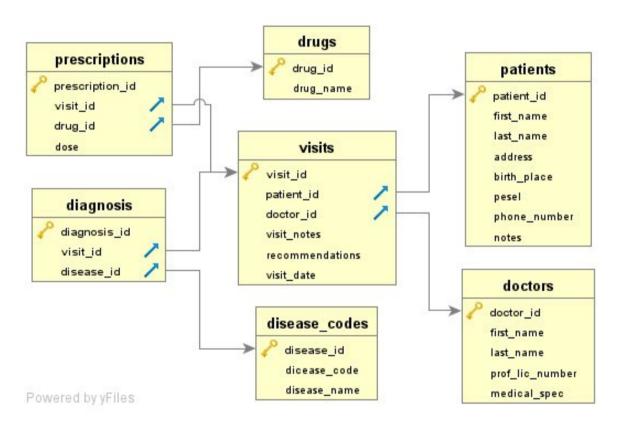


Rys. 1 - Schemat projektu bazy danych

4. Implementacja projektu

Baza danych została zaimplementowana przy pomocy silnika PostgreSQL. Postgres jest potężnym systemem obiektowo-relacyjnych baz danych typu open source, aktywnie rozwijanym od ponad 30 lat [5].

Baza przedstawia się następująco:



Rys. 2 - Schemat zaimplementowanej bazy danych. [Obraz uzyskany przy pomocy narzędzia DbVisualizer (Free)]

Względem projektu bazy pewne kolumny zostały zmienione/zredukowane w celu optymalizacji wynikowej bazy danych.

Zaimplementowana baza danych jest przykładowym zbiorem relacji danych w przychodni. Uwzględnieni zostali pacjenci oraz lekarze, jako uproszczenie grupy pracowników danej placówki. Pacjent może mieć wiele wizyt a owa wiele towarzyszących jej aspektów jak np.: recepta i diagnoza.

❖ Kod SQL:

CREATE DATABASE **cwmbd_project**

WITH OWNER = postgres

ENCODING = 'UTF8'

LC_COLLATE = 'Polish_Poland.1250'

LC_CTYPE = 'Polish_Poland.1250'

TABLESPACE = pg_default

CREATE TABLE patients(

CONNECTION LIMIT = -1;

patient_id SERIAL PRIMARY KEY,

first_name VARCHAR(20),

last_name VARCHAR(20),

address VARCHAR(100),

birth_place VARCHAR(20),

pesel VARCHAR(11) UNIQUE,

phone_number VARCHAR(20),

notes VARCHAR(100));

CREATE TABLE **drugs**(

drug_id SERIAL PRIMARY KEY,
 drug_name VARCHAR(20));

CREATE TABLE disease_codes(

disease_id SERIAL PRIMARY KEY,
dicease_code VARCHAR(5) UNIQUE,
 disease name VARCHAR(100));

CREATE TABLE **doctors**(

doctor_id SERIAL PRIMARY KEY,

first_name VARCHAR(20),

last_name VARCHAR(20),

prof_lic_number VARCHAR(7) UNIQUE,

medical_spec VARCHAR(20));

CREATE TABLE **visits**(

visit_id SERIAL PRIMARY KEY,
patient_id SERIAL REFERENCES patients,
doctor_id SERIAL REFERENCES doctors,
visit_notes VARCHAR(200),
recommendations VARCHAR(200),
visit_date DATE);

CREATE TABLE prescriptions(

CREATE TABLE diagnosis(

diagnosis_id SERIAL PRIMARY KEY, visit_id SERIAL REFERENCES visits, disease_id SERIAL REFERENCES disease_codes);

4.1. Wdrożenie rozwiązania

Baza została wdrożona na bezpłatny serwer przy pomocy serwisu Heroku. Przed utworzeniem obiektu bazy w serwisie Heroku należało stworzyć umownie obiekt aplikacji. Następnie zainstalować dodatek zapewniający silnik PostgreSQL dla owej aplikacji. Panel właściciela aplikacji Heroku umożliwia podgląd bazy dla zapytań tylko do odczytu oraz poznanie parametrów dostępu do bazy danych. Parametry te ulegają częstym zmianom. Oprócz standardowego połączenia z bazą danych przy pomocy adresu hosta, nazwy bazy danych, użytkownika i hasła, Heroku umożliwia dostęp do aplikacji/bazy danych za pośrednictwem narzędzia linii poleceń Heroku CLI.

Bazy danych hostowane za pośrednictwem serwisu Heroku przechowywane są w formie zaszyfrowanego woluminu, w ten sam sposób, przez cały okres jego życia przez usługę Amazon. Baza danych nie jest szyfrowana z poziomu silnika bazy danych. Taką funkcję należy zaimplementować samemu poprzez np.: skorzystanie z gotowych bibliotek.

Szyfrowanie z poziomu SZBD wymaga zaufania administratorowi bazy danych ze względu na to, że zakodowane wyjścia i klucze przechowywane są na tym samym serwerze. Inną opcją zabezpieczenia danych jest szyfrowanie z poziomu aplikacji klienckiej, w ten sposób zabezpiecza się hasła do kont poszczególnych użytkowników danego serwisu. W przypadku zapomnienia hasła, nie można uzyskać przypomnienia, hasło można jedynie zresetować.

5. Zabezpieczenia bazy danych

Baza danych PostgreSQL hostowana za pośrednictwem serwisu Heroku posiada kilka gotowych zabezpieczeń zarówno typowych dla tego silnika baz danych jak i platformy wdrażającej,

5.1. Szyfrowanie danych

Jak już wspomniano, cała baza danych wdrożona na serwerze amazona jest zaszyfrowana. Dodatkowym zabezpieczeniem może być szyfrowanie z poziomu SZBD przy pomocy modułu pgcrypto, zapewniającego funkcje kryptografii [6].

Rys. 3 - Terminal po zalogowaniu się za pośrednictwem narzędzia linii poleceń Heroku CLI z komenda psql - utworzenie rozszerzenia silnika oraz sprawdzenie działania.

Dane, które możemy chcieć zabezpieczyć podlegają pewnym praktykom, adekwatnym do istniejącego/zakładanego zagrożenia. W przypadku danych medycznych, należy obawiać się o dane osobowe i powiązane, uznane za wrażliwe w myśl istniejących przepisów - np.: wszelkie dane z wizyt lekarskich, przebieg chorób.

W zależności od przyjętej strategii można wybrać różne podejścia do szyfrowania danych, w tym:

❖ Zabezpieczenie bazy danych szyfrując cały obszar pamięci jaki zajmuje z poziomu systemu operacyjnego - poszczególne dane z tabel są na bieżąco odszyfrowywane podczas stosowania zapytań do bazy,

- ❖ Szyfrowanie poszczególnych tabel i kolumn ukrywanie treści przez administratora lub osobę uprawnioną, która posiada odpowiednie klucze, potrzebne do odczytania prawdziwej treści.
- Szyfrowanie haseł, wartości które nie podlegają odczytowi, a weryfikacji, tylko osoba, która tą treść zna i dostarczy ją do systemu, uzyska dostęp. Treść ta znajduje się w postaci zaszyfrowanej w bazie danych, nigdy nie inaczej. Po podaniu w aplikacji klienckiej/panelu administratora, jest szyfrowana i system sprawdza, czy w bazie znajduje się odpowiadający temu rekord.

Szyfrowanie danych całego woluminu pamięci zostało już zrealizowane przez usługę serwerowni firmy Amazon. Firma ta posiada zabezpieczenia trudne do zweryfikowania dla klienta, więc w ramach projektu wykonane zostaną dwie pozostałe metody.

5.2. Szyfrogramy numeru pesel

W tym podejściu, numery pesel zostaną przeniesione do oddzielnej tabeli. W przychodni będzie potrzebny numer pesel aby zidentyfikować pacjenta, numer ID zaszyfrowanej wersji peselu będzie prowadził jako klucz obcy do rekordów tego pacjenta w innych tabelach.

atient_id	first_name	last_name	address	birth_place	phone_number	notes	pesel
1	Kamil	Suchenek	B awatkowa 9/13	Gliwice	467598762		467598762
2	Dominika	Gab r	Nak o îl skie Morcinka 9	Tarnowskie G ry	765398267	al.tetracykliny	765398267
3	Agnieszka	Rudziun	Tarnogorska 8 Stare Tarnowice	Oťwiŕcim	654329873	-	654329873
4	Natalia	Matuszek	S oneczna 5 Krupski M yn	Lubliniec	876539846	agorafobia	876539846
5	Marcin	Kasprzyk	Celiny 8 Tapkowice	Ruda îl-ska	987654323	po odwyku	987654323
ojekt-cyber ojekt-cyber	::DATABASE=> ::DATABASE->						
ojekt-cyber ojekt-cyber TER TABLE ojekt-cyber		DROP COLUMN select*from	pesel; patients;	birth_place	phone_number	notes	
ojekt-cyber TER TABLE ojekt-cyber	r::DATABASE-> r::DATABASE=> first_name	DROP COLUMN select*from last_name +	pesel; patients; address	· 	.	notes	
ojekt-cyber ojekt-cyber TER TABLE ojekt-cyber	r::DATABASE-> r::DATABASE=> first_name Kamil	DROP COLUMN select*from last_name + Suchenek	pesel; patients; address		467598762	 	
ojekt-cyber ojekt-cyber TER TABLE ojekt-cyber atient_id	r::DATABASE-> r::DATABASE=> first_name	DROP COLUMN select*from last_name +	pesel; patients; address	· 	.	notes 	
ojekt-cyber ojekt-cyber TER TABLE ojekt-cyber atient_id	c::DATABASE-> c::DATABASE=> first_name	DROP COLUMN select*from last_name Suchenek Gab r	pesel; patients; address	Gliwice Garnowskie G~ry	467598762 765398267	 	

Rys. 4 - Fragment terminala - usunięcie kolumny z numerami pesel.

Utworzenie tabeli peseli i umieszczenie w niej nowych zaszyfrowanych numerów:

```
projekt-cyber::DATABASE=> CREATE TABLE pesels(
projekt-cyber::DATABASE(> pesel id SERIAL PRIMARY KEY,
projekt-cyber::DATABASE(> pesel VARCHAR UNIQUE);
CREATE TABLE
projekt-cyber::DATABASE=> INSERT INTO pesels (pesel) VALUES
projekt-cyber::DATABASE-> (crypt('11111111111',GEN_SALT('md5'))),
projekt-cyber::DATABASE-> (crypt('2222222222',GEN_SALT('md5'))),
projekt-cyber::DATABASE-> (crypt('3333333333',GEN_SALT('md5'))),
projekt-cyber::DATABASE-> (crypt('44444444444',GEN_SALT('md5'))),
projekt-cyber::DATABASE-> (crypt('5555555555',GEN SALT('md5')));
INSERT 0 5
projekt-cyber::DATABASE=> SELECT * FROM pesels;
 pesel id
                                  pesel
          1 | $1$1gJBcscr$ucCciRe7mnSotjVbB7mCk1
           2 | $1$4hPsCpKf$VdV1B2fytWP2LOpKG1/aC1
           3 | $1$uN01sugz$VZaw7P8s1X4ZXGDz6AK8L1
          4 | $1$C5KTRWU7$F0TtMXLOaVfN9uRzEFVrp0
           5 | $1$i9Qqix15$NUPg6hK7BIL.z8w5viDrx0
(5 wierszy)
```

Rys. 5 - Fragment terminala - utworzenie tabeli peseli oraz wstawienie nowych wartości.

Wartości wstawione nie posiadają znanej wady szyfrowania md5, zaszyfrowanie nawet 2 takich samych wartości zwróci 2 inne, dzięki dodaniu "soli" do algorytmu, utrudnia to również odnalezienie oryginalnego tekstu i komplikuje wynik. Aktualnie algorytm md5 uchodzi za niezbyt bezpieczne rozwiązanie, poleca się zastosowanie w zamian algorytmów SHA-N albo dodanie soli.

❖ Zapytanie z poziomu klienta:

Po dodaniu kolumny:

ALTER TABLE patients ADD COLUMN pesel_id SERIAL UNIQUE REFERENCES pesels;

oraz wstawieniu wartości pesel id do tabeli patients:

```
UPDATE patients SET pesel_id = 1 WHERE patient_id = 1;
UPDATE patients SET pesel_id = 2 WHERE patient_id = 2;
UPDATE patients SET pesel_id = 3 WHERE patient_id = 3;
UPDATE patients SET pesel_id = 4 WHERE patient_id = 4;
UPDATE patients SET pesel_id = 5 WHERE patient_id = 5;
```

Rys. 6 - Fragment terminala - odszukanie rekordu pacjenta.

Tego rodzaju zapytanie odszukuje rekord pacjenta o ile poda się jego pesel.

❖ Porównanie czasu zapytania z szyfrowaniem i bez:

Zestawiono czas realizacji zapytania na serwerze lokalnym, przy pierwszym uruchomieniu zapytania.

bez szyfrowania: 193ms aż do 112-125ms przy powtarzaniu operacji:

> z szyfrowaniem: 173ms aż do 130-150ms przy powtarzaniu operacji:

Zapytanie bez szyfrowania częściej zwraca szybciej wynik niż zapytanie z szyfrowaniem, różnica jest niewielka, gdyż dotyczy szyfrowania tylko jednej wartości a nie ich zbioru.

5.3. Szyfrowanie dwukierunkowe

Metoda ta jest odpowiednia, gdy mamy zamiar zataić zawartość przed niepożądanymi podmiotami, a udostępnić ją zweryfikowanym użytkownikom.

Rozszerzenie pgcrypto oferuje funkcje szyfrowania PGP. Obrana metoda wykorzystuje dwa klucze: publiczny, który służy do zaszyfrowania treści oraz prywatny, dzięki któremu można odczytać treść oryginalną. W celu wygenerowania kluczy wykorzystano narzędzie Gnupgp (https://www.gnupg.org/download/).

Procedura wygląda następująco:

modyfikacja bazy danych - zmiana typów danych na bytea, w celu przechowywania postaci zaszyfrowanej:

```
CREATE TABLE patients(
    patient_id SERIAL PRIMARY KEY,
    first_name bytea,
    last_name bytea,
    address bytea,
    birth_place bytea,
    pesel bytea UNIQUE,
    phone_number bytea,
    notes bytea);
```

utworzenie pary kluczy przy pomocy oprogramowania Gnupgp,

Klucze utworzono i wyeksportowano przy pomocy poleceń poleceń zaznaczonych na żółto z rysunku nr 7. Polecenie edit-key pozwala na odczytanie identyfikatorów kluczy, które są potrzebne w celu ich eksportu do plików.

F.26.3.9. Generating PGP Keys With GnuPG

-a from the command.

To generate a new key:

| gpg --gen-key |
| The preferred key type is "DSA and Elgamal".
| For RSA encryption you must create either DSA or RSA sign-only key as master and then add an RSA encryption subkey with gpg --edit-key.
| To list keys:
| gpg --list-secret-keys |
| To export a public key in ASCII-armor format:
| gpg -a --export KEYID > public key |
| To export a secret key in ASCII-armor format:
| gpg -a --export-secret-keys KEYID > secret key |

Rys. 7 - Fragment dokumentacji pgcrypto [odnośnik bez numeru 2]

You need to use dearmor() on these keys before giving them to the PGP functions. Or if you can handle binary data, you can drop

❖ zaimplementowanie szyfrowania w poleceniu INSERT:

Rys. 8 - Instrukcja wsadowa dla szyfrowania danych kluczem publicznym.

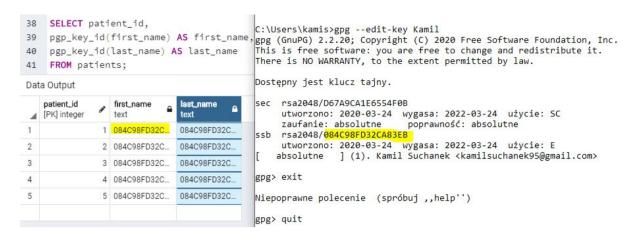
Na rysunku nr 8, przedstawiono wstawianie danych do tablicy patients, w miejscach 'key' znajduje się interpretacja klucza w ASCII:

----BEGIN PGP PUBLIC KEY BLOCK---- moenbefsblubcadnoxslbt/ougrmtinckplygmjvclbvf8hc8suew,/Zil28s5Pjo d7aR/kvk/bkBD96/06yW0v99Yf3zlqoZUgylOqnpBC100TmWPYTFcb50KlnDwa D85d70PG9EHvlAsGpAKDWIZhGiQd01+TUztb2OjBFxf78w2tkBc47O4lqRGHMkhl iJGxoCapaVFGyuT6RR+/IDOm20k7rsFNgVm4q/iF5r0j8xZkkph0oxD++CgCYjwK eqdYScVcITO7OLW+iRsOADBp2]0rsj9w2FEreWFoFHZhateV+K5EnfzCWPGFuR1 KmvAcDycoopM5f9ygeHa7+bHHvhcXckj2qvAbeBAAG0kkthbWisiFN1Y2hhbmVr IDxrYW1pbHN1ZyhhbmVr0TVAZ21JAWwuY29F0kBVAOTAOQAPhYhBj/cTk03Sw6j 30+10NZ6nKHmVUBLBQjeemy1AhsDB0kDwmcABQsjCAcCBhUKCQgLAqQWAgMBAh4B AheAAAQjeNZ6nKHmVUBLGQolAMIHjr/12gWHFfS0/10xLUIUMOwF6H0kSGmhP2V SjDWer2Y+g4C6pqK8hMKX1FsQf8m/jBX+cBWbYuPDjlgPb3UiS8Fk5uASLxMe17a /EPyW/SpgPlPvrjaTedGqbkjp1Z5zn79ywGBYBqdqsN2qD07Z0q3hFqfMXEPtxp8 Fm4VnsnMkHytGjihZDVMBL+Em+C0a0Vyhcc5p-7pdy/sctyEVCX7mm+LHig7HvW, v1THLxhall2Wgse0qWtFxFkZk17yxjRjDkcjG38jUGSj9/zsq/9xzYTF4QMJX4IR /Y513Wsox5Qp5llrxA4h-lboUssvr9+NWmduhn3trVsovUpOsAQ0EXnpst0eIhAM13 ugja07LivZixPlyrsRtd10e0TB2755f1YWdtHUHlj9F50kD2g7lxkHXjWD+gG5Qb Adm9PWS/kxt+l/yvapMF/6Whe7j/qG55Pav+51YeAA73PorH6L0kD/WcHKP57yxw ZUW6iadwkahqLePYEq2ekLkdYhwMmDx4fluHtGW0yfTAzEOWMjkl8H116kECzmy aeYfrPumylxXQXdWwa+Y5Cgq2R8hjDXTTXTTHEww4kvkcpzNf8w5jy7ocRmx5vQ sdWX7zE7kHvQjC78kvb1s314ld8ccb6rd3amNs+nhf0H91A5slAUWmrYGuQH0uewWW TSY7RsvbhbcADcVrlwUAEQEAAYkBPAQYAQgAJhYnBj/cTkD3Sw6j30-110NZ6nKHm vU8lBQjeemy1AnsMBQkDwmcAAAOJENZ6nKHmWUBLA1QHX7jryxkR6SaduREMre xbe17cdbAFrbjMHxQA9XEeif2mtUV+qAVBRGneikL+oAkbnyulq\b2koKvD+icfP AYWCMd9mMeAwxbqyCa3dfcTb2aEbeV2tjTg7zg265Nd4tPk2zpbcZglyJuA9001Ac MkXspC03+QmwmfEd2Y4wX5AmcQ6fq1U4TkQiKWVI+zDr8EXYAGPKp0cV/AetjU/o E040Rr3U0rRDyef610VKwBqpu83bzr4U56kngM1cRTd16oTLJCzudnifTzUHyFGU VTTpLsnfyvwKWyNFpG3XXmM/QrFTYsk/f4/C60dW6eimjNTyjsY1wiCcAwxVMH gZM= pLj0 -----ENDPGPPUBLIC KEY BLOCK-----

sprawdzenie rozwiązania:

Dat	Data Output							
4	patient_id [PK] integer	first_name bytea	last_name bytea	address bytea	birth_place bytea	pesel bytea		
1	1	[binary data]	[binary data]	[binary data]	[binary data]	[binary data]		
2	2	[binary data]	[binary data]	[binary data]	[binary data]	[binary data]		
3	3	[binary data]	[binary data]	[binary data]	[binary data]	[binary data]		
4	4	[binary data]	[binary data]	[binary data]	[binary data]	[binary data]		
5	5	[binary data]	[binary data]	[binary data]	[binary data]	[binary data]		

Rys. 9 - fragment narzędzia PgAdmin4 - dane po zaszyfrowaniu.



Rys. 10 - Fragment narzędzia PgAdmin4 oraz terminala (po prawej) - kolorem żółtym zaznaczono pokrywające się numery ID zastosowanego klucza.

43 SELECT patient id.

7-	orrect bac	3 TO 3 T - 3 TO 2							
44	pgp_pub_de	pgp_pub_decrypt(first_name, keys.privkey, 'password') AS first_name,							
45	pgp_pub_de	pgp_pub_decrypt(last_name, keys.privkey, 'password') AS last_name,							
46	pgp_pub_de	crypt(pesel, keys	s.privkey, 'passwo	rd') AS pesel					
47	FROM patients								
48	CROSS JOIN								
49	(SELEC	(SELECT dearmor('BEGIN PGP PRIVATE KEY BLOCK							
50									
51	lQPGBF56bL	UBCADNoxSLBT/ou9F	RMtlNcXbJVgmjvclDV	f8NG8Suew/XZI2Bs5Pj0					
52	d7aR/kvk/Dl	kBRD96/Q6yW0vt90\	f32tQ0ZUgyl0qnpBC	lo0TmWPYTFcb50KJNDwa					
L.3		DITAGEN KOM TACTO	1401 TH1+61070Fv.f7	9mO+K9C4Z04TaDCUMkki					
Dai	ta Output								
4	patient_id [PK] integer	first_name text	last_name text □	pesel text					
⊿ 1									
1 2	[PK] integer 1	text	text	text					
8.	[PK] integer 1	text Kamil	text Suchenek	text 43898765672					
2	[PK] integer 1 2	text Kamil Dominika	text Suchenek Gabór	text 43898765672 87654678926					
2	[PK] integer 1 2	text Kamil Dominika Agnieszka Natalia	text Suchenek Gabór Rudziun	text 43898765672 87654678926 67489356728					

Rys. 11 - Fragment narzędzia PgAdmin4 - Proces odszyfrowywania danych.

Uzyskiwanie danych wymaga klucza prywatnego oraz hasła, które można dodać przy generowaniu klucza.

- Zestawienie czasu realizacji zapytania z szyfrowaniem i bez:
 - bez szyfrowania: 133ms aż do 119-130ms powtarzając czynność:

SELECT patient id, first name, last name, pesel FROM patients: select*from patients:

> z szyfrowaniem: 1 sekunda i 459ms aż do sekundy i 440ms:

SELECT patient id. pgp pub decrypt(first name, keys.privkey, 'passoword ') AS first name, pgp pub decrypt(last name, keys.privkey, 'passoword ') AS last name, pgp pub decrypt(pesel, keys.privkey, 'passoword ') AS pesel FROM patients **CROSS IOIN** (SELECT dearmor('----BEGIN PGP PRIVATE KEY BLOCK----

IQPGBF56bLUBCADNoxSLBT/ou9RMtINcXbJVgmjvcIDVf8NG8Suew/XZI2Bs5PjO =79ID
----END PGP PRIVATE KEY BLOCK----

') As privkey) As keys;

Tym razem różnica w czasie realizacji zapytania jest znaczna.

5.4. Szyfrowanie symetryczne

Wykonano szyfrowanie jednym kluczem, służącym zarówno do ukrywania jak i wyjawiania zaszyfrowanej treści. Klucz taki często zawarty jest po stronie klienta, w ten sposób dane nie są zagrożone od strony administratora bazy danych.

Klucz może stanowić prawie dowolny ciąg znaków. zastosowano ponownie funkcje oferowane przez rozszerzenie pącrypto. Kod:

```
INSERT INTO
patients(first_name,last_name,address,birth_place,pesel,phone_number,notes)
SELECT pgp_sym_encrypt(pat.fn, 'password') AS first_name,
    pgp_sym_encrypt(pat.ln, 'password') AS last_name,
    pgp_sym_encrypt(pat.ad, 'password') AS address,
    pgp_sym_encrypt(pat.bi, 'password') AS birth_place,
    pgp_sym_encrypt(pat.pe, 'password') AS pesel,
    pgp_sym_encrypt(pat.ph, 'password') AS phone_number,
    pgp_sym_encrypt(pat.no, 'password') AS notes
FROM (VALUES
('Kamil','Suchenek','Bławatkowa 9/13','Gliwice','43898765672','467598762','---'),
('Dominika','Gabór','Nakło Śląskie Morcinka 9','Tarnowskie Góry','87654678926','765398267','al.tetracykliny'),
('Agnieszka','Rudziun','Tarnogorska 8 Stare Tarnowice','Oświęcim','67489356728','654329873','-'),
('Natalia','Matuszek','Słoneczna 5 Krupski Młyn','Lubliniec','78754378961','876539846','agorafobia'),
('Marcin','Kasprzyk','Celiny 8 Tąpkowice','Ruda Śląska','87546783218','987654323','po odwyku') )
AS pat(fn,ln,ad,bi,pe,ph,no);
```

❖ Zaszyfrowane dane:

32									
Data Output									
4	patient_id [PK] integer	first_name bytea	last_name bytea	address bytea	birth_place bytea	pesel bytea	phone_number bytea	ø	
1	1	[binary data]	[binary data]	[binary data]	[binary data]	[binary data	[binary data]		
2	2	[binary data]	[binary data]	[binary data]	[binary data]	[binary data	[binary data]		
3	3	[binary data]	[binary data]	[binary data]	[binary data]	[binary data	[binary data]		
4	4	[binary data]	[binary data]	[binary data]	[binary data]	[binary data	[binary data]		
5	5	[binary data]	[binary data]	[binary data]	[binary data]	[binary data	[binary data]		

Rys. 12 - Fragment narzędzia PgAdmin4 - Zaszyfrowane dane.

❖ Zapytanie klienta do odszyfrowania danych:

```
SELECT patient_id,
pgp_sym_decrypt(first_name, 'password') AS first_name,
pgp_sym_decrypt(last_name, 'password') AS last_name,
pgp_sym_decrypt(pesel, 'password') AS pesel
FROM patients;
```

- ❖ Zestawienie czasu zapytań z szyfrowaniem i bez:
 - > z szyfrowaniem: 150ms aż do 140ms

SELECT patient_id, pgp_sym_decrypt(first_name, 'password') AS first_name, pgp_sym_decrypt(last_name, 'password') AS last_name, pgp_sym_decrypt(pesel, 'password') AS pesel FROM patients;

> bez szyfrowania: 133ms aż do 119-130ms powtarzając czynność:

SELECT patient_id, first_name, last_name, pesel FROM patients; select*from patients;

Podobnie jak w przypadku szyfrowania peselu, różnica w czasie realizacji zapytań z szyfrowaniem i bez istnieje, lecz jest niewielka.

6. Podsumowanie

Zaprojektowano, zaimplementowano i wdrożono przykładową, wzorcową bazę danych medycznych, dla danego aspektu pracy przychodni. Przedstawiono i zastosowano dwa sposoby zabezpieczenia bazy przy pomocy szyfrowania danych za pomocą algorytmu md5 z pomocą "soli" oraz pary kluczy szyfrowania dwukierunkowego.

Opracowanie stanowi podstawowe, zręczne wprowadzenie do praktycznego aspektu szyfrowania danych stosując silnik bazodanowy PostgreSQL.

Literatura:

- [1] "SQL Wydanie III" Danuta Mendrala, Marcin Szeliga
- [2] "Php i MySQL Od nowicjusza do wojownika ninja" Kevin Yank
- [3] Wykłady Systemy informatyczne w medycynie "Zarządzanie indywidualną dokumentacją medyczną" Piotr Zarychta
- [4] postgresql.org/docs/9.1/auth-pg-hb-a-conf.html [dostęp 27.03.2020]
- [5] postgresql.org/[dostep 14.03.2020]
- [6] postgresql.org/docs/current/pgcrypto.html [dostęp 24.03.2020 "current" oznacza wersję PostgreSQL 12].

<u>Dodatkowe źródła merytoryczne:</u>

- stackoverflow.com/questions/24496536/best-practices-for-using-pgcrypto-pgpencryption-with-heroku-and-rails
- stackoverflow.com/questions/10752456/how-to-enable-contrib-modules-on-heroku-postgres-database
- postgresonline.com/journal/archives/165-Encrypting-data-with-pgcrypto.html
- severalnines.com/database-blog/how-secure-your-postgresgl-database-10-tips
- dbrnd.com/2016/03/postgresql-best-way-for-password-encryption-usingpgcryptos-cryptographic-functions/

[dostęp do powyższych źródeł 26.03.2020]

Zbiór skryptów jest dostępny na repozytorium: github.com/KamilSuchanek95/projekt-cyber

[Gałąź master zawiera podstawową implementację bazy bez szyfrowania, gałąź pesel zawiera implementację dla szyfrowania tylko numeru pesel, a gałąź gpg zawiera skrypty dla szyfrowania całej tabeli pacjentów, kod dla szyfrowania symetrycznego nie został uwzględniony, jest analogiczny dla gałęzi gpg z drobną zmianą ujętą w przedstawionym sprawozdaniu.]