



КООРДИНАЦИОННЫЙ ЦЕНТР  
ДОМЕНОВ .RU/.РФ

# Угрозы визуальной схожести в доменной и почтовой адресации

Вадим Михайлов  
Консультант по инфраструктуре  
Вебинар  
09 июля 2024

cctld.ru | кц.рф

.RU

.РФ



## Определения

### RFC 6365: Terminology Used in Internationalization in the IETF

#### **Язык**

Способ коммуникации людей

#### **Скрипт**

Набор графических символов, используемых для написания одного или более языков

#### **Письменность**

Набор правил использования одного или более скриптов, чтобы писать на отдельно взятом языке

#### **Символ**

Минимальная единица письменности

#### **Глиф**

Графическое изображение символа

## Омоглифы

Омоглиф – от древнегреческого ὁμός – «одинаковый», и γλυφή – «знак»

## Омоглифы

Омоглиф – от древнегреческого ὁμός – «одинаковый», и γλυφή – «знак»

– латинская «о» (U+006F), кириллическая «о» (U+043E) и греческая «о» (U+03BF)

## Омоглифы

Омоглиф – от древнегреческого ὁμός – «одинаковый», и γλυφή – «знак»

- латинская «о» (U+006F), кириллическая «о» (U+043E) и греческая «о» (U+03BF)
- расширенная латиница а (U+0061) – ä (U+00E4)

## Омоглифы

Омоглиф – от древнегреческого ὁμός – «одинаковый», и γλυφή – «знак»

- латинская «о» (U+006F), кириллическая «о» (U+043E) и греческая «о» (U+03BF)
- расширенная латиница а (U+0061) – ä (U+00E4)
- цифра «0» (U+0030) и заглавная буква «O» (U+004F)

## Омоглифы

Омоглиф – от древнегреческого ὁμός – «одинаковый», и γλυφή – «знак»

- латинская «o» (U+006F), кириллическая «о» (U+043E) и греческая «ο» (U+03BF)
- расширенная латиница а (U+0061) – ä (U+00E4)
- цифра «0» (U+0030) и заглавная буква «O» (U+004F)
- строчная m (U+006D) и рядом rn (U+0072 и U+006E)

## Лигатура

Лигатура – от латинского *ligature* – «связка»

- латинская «æ» (U+00E6) сочетает буквы «a» (U+0061) и «e» (U+0065)
- латинская «ffl» (U+FB04) является соединением букв «f» (U+0066), «f» (U+0066) и «l» (U+006C)
- кириллическая «лъ» (U+0459) представляет собой соединение букв «л» (U+043B) и «ь» (U+044C)

## Омографы

Омограф – от древнегреческого ὁμός – «одинаковый», и  
γράφω – «писать»

- на латинице «*tata*» (U+006D U+0061 U+006D U+0061) и  
на кириллице «*tata*» (U+0442 U+0430 U+0442 U+0430)
- на латинице «*rau*» (U+0070 U+0061 U+0079) и на  
кириллице «*rau*» (U+0440 U+0430 U+0443)

## Паронимы

Паро́нимы – от древнегреческого *παρά* — «около; рядом», и *ὄνυμα* — «имя»)

- адресат и адресант
- экскаватор и эскалатор
- live и leave
- Кастанаевская и Кустанайская

Создание вредоносных ресурсов и привязка к ним доменных имен максимально похожих на легитимные

При помощи омоглифов:

- gosuslugi.ru и gosus<sup>1</sup>ugi.ru
- paypal.com и paypa<sup>1</sup>.com
- bank.com и b<sup>ä</sup>nk.com/ba<sup>ñ</sup>k.com/ba<sup>ñ</sup>k.com
- госуслуги.рф и госуслу<sup>ти</sup>.рф

Создание вредоносных ресурсов и привязка к ним доменных имен максимально похожих на легитимные

При помощи омографов:

- pay.com (Latin) и **pay**.com (Кириллица)
- cop.online (Latin) и **cop**.online (Кириллица)
- *mama.shop* (Latin) и ***mama***.shop (Кириллица)

Создание вредоносных ресурсов и привязка к ним доменных имен максимально похожих на легитимные

При помощи лигатур:

- crownbank.org и cro**vv**nbank.org
- raiffeisen.com и rai**ff**eisen.com
- стальпром.ею и ста**ль**пром.ею

Создание вредоносных ресурсов и привязка к ним доменных имен максимально похожих на легитимные

## Typosquatting:

– bank.ru и bark.ru

## Combosquatting:

– авито.рф и авито-услуги.рф

## TLDsquatting:

– bank.com и bank.co

## Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Сценарии

– Взломанный реальный контакт

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Сценарии

- Взломанный реальный контакт
- Фиктивный доверенный контакт

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Сценарии

- Взломанный реальный контакт
- Фиктивный доверенный контакт
- Фиктивный хорошо известный контакт

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Реализация

— Взлом почты

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Реализация

- Взлом почты
- Использование уязвимостей клиента

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Реализация

- Взлом почты
- Использование уязвимостей клиента
- Подделка заголовков писем

# Business Email Compromise

Business Email Compromise – многоступенчатые атаки связанные с компрометацией корпоративной переписки

## Реализация

- Взлом почты
- Использование уязвимостей клиента
- Подделка заголовков писем
- Использование схожих адресов

# Business Email Compromise

**some@example.com**

# Business Email Compromise

**some@example.com**

## Business Email Compromise

some@example.com

some@exa**rn**ple.com

## Business Email Compromise

some@example.com

some@exa**rn**ple.com

so**rne**@example.com

## Business Email Compromise

some@example.com

some@exa**rn**ple.com

so**rne**@example.com

so**o**me@example.com

## Business Email Compromise

some@example.com

some@exa**rn**ple.com

so**rne**@example.com

so**o**me@example.com

## ФБР Internet Crime Report 2022-2023



**\$12.5  
МЛРД**

ВЕС – \$2 946 830 270

Фишинг – \$18 728 550

## ФБР Internet Crime Report 2022-2023

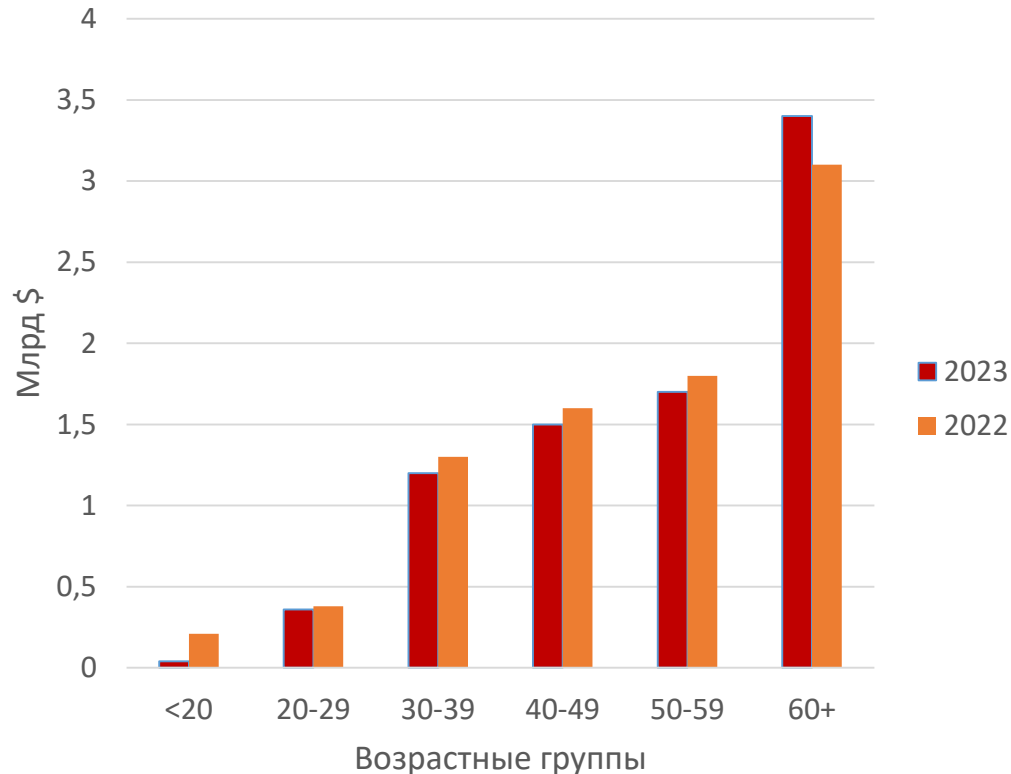


**\$12.5  
МЛРД**

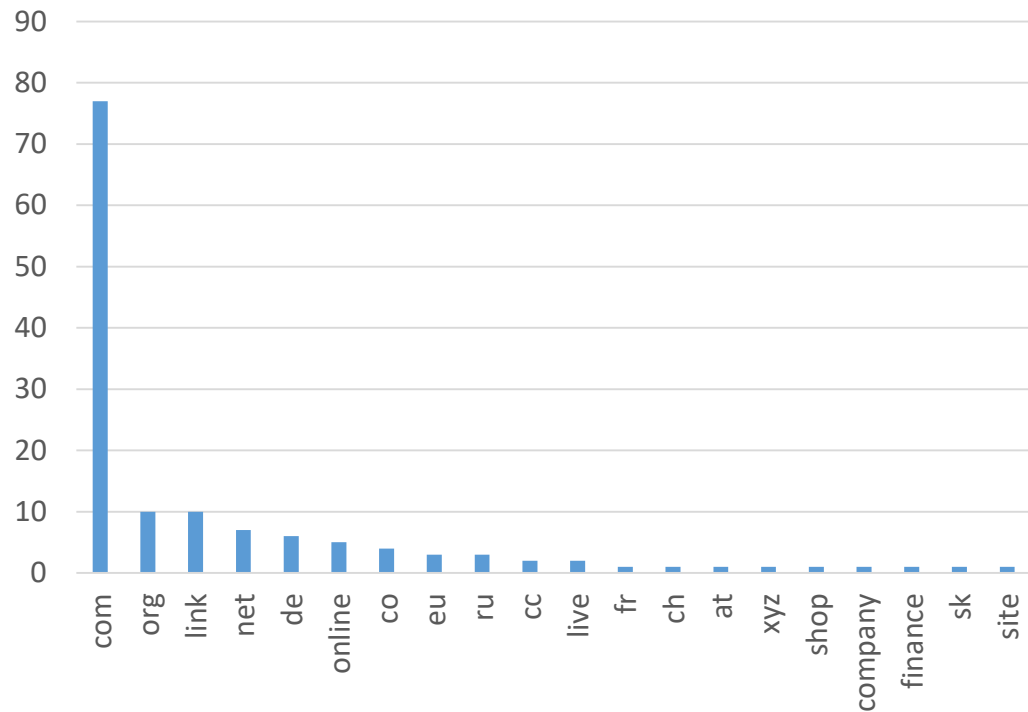
ВЕС – \$2 946 830 270

Фишинг – \$18 728 550

## Статистика



## Статистика



| Domain Extension | Percentage (%) |
|------------------|----------------|
| com              | 77             |
| org              | 10             |
| link             | 10             |
| net              | 7              |
| de               | 6              |
| online           | 5              |
| co               | 4              |
| eu               | 3              |
| ru               | 3              |
| cc               | 2              |
| live             | 2              |
| fr               | 1              |
| ch               | 1              |
| at               | 1              |
| xyz              | 1              |
| shop             | 1              |
| company          | 1              |
| finance          | 1              |
| sk               | 1              |
| site             | 1              |

## Кейсы

— Фишинговый ресурс компании Adobe

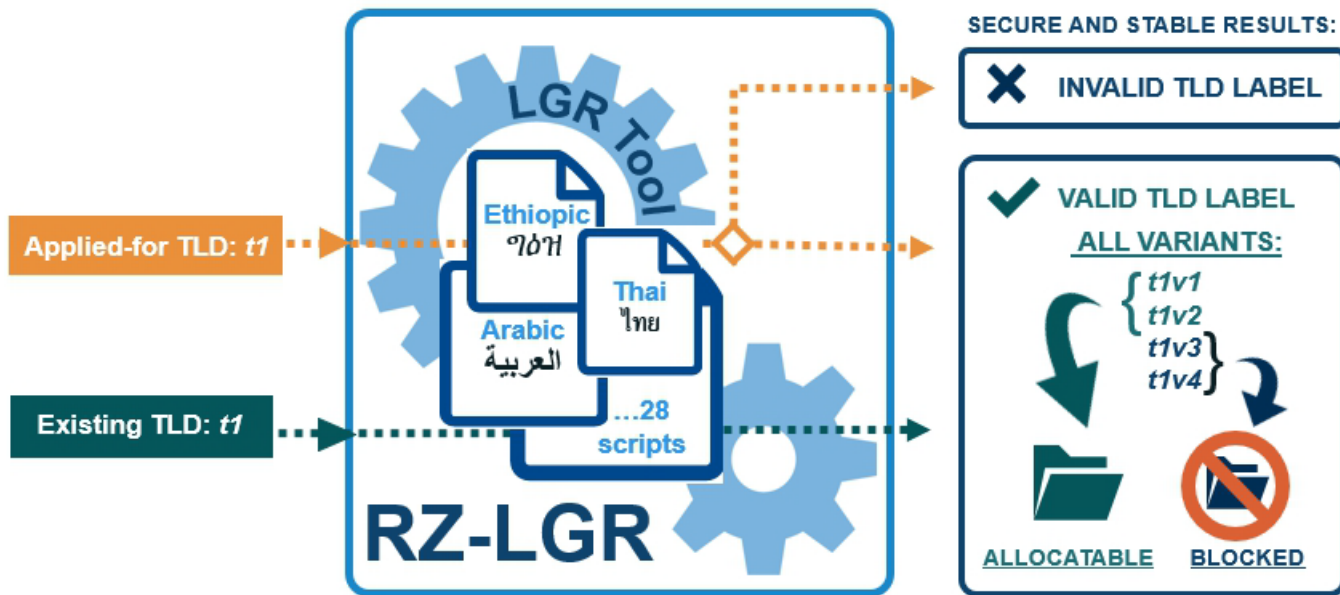
ado**b**e.com – распространение трояна BetaBot

## Кейсы

- Фишинговый ресурс компании Adobe  
ado**b**e.com – распространение трояна BetaBot
- Уязвимость в ПО Apple CVE-2018-4277  
«d<sub>z</sub>» (U+A771) отображалась как «d» (U+0064)

## Решения

— На уровне ICANN – Label Generation Rules



## Решения

- На уровне ICANN – Label Generation Rules
- На уровне Регистратур – запрет кросскрипта и IDN-таблицы



bank.com

.РФ – только русский алфавит

.RU – только базовая латиница



bank.плата.com

## Решения

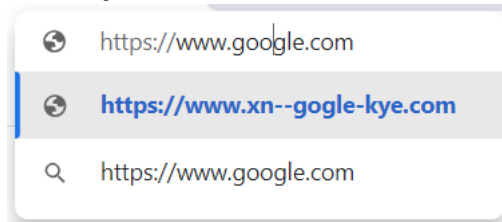
- На уровне ICANN – Label Generation Rules
- На уровне Регистратур – запрет кросскрипта и IDN-таблицы
- На уровне Регистраторов – homoglyph bundling

для **bank.com**

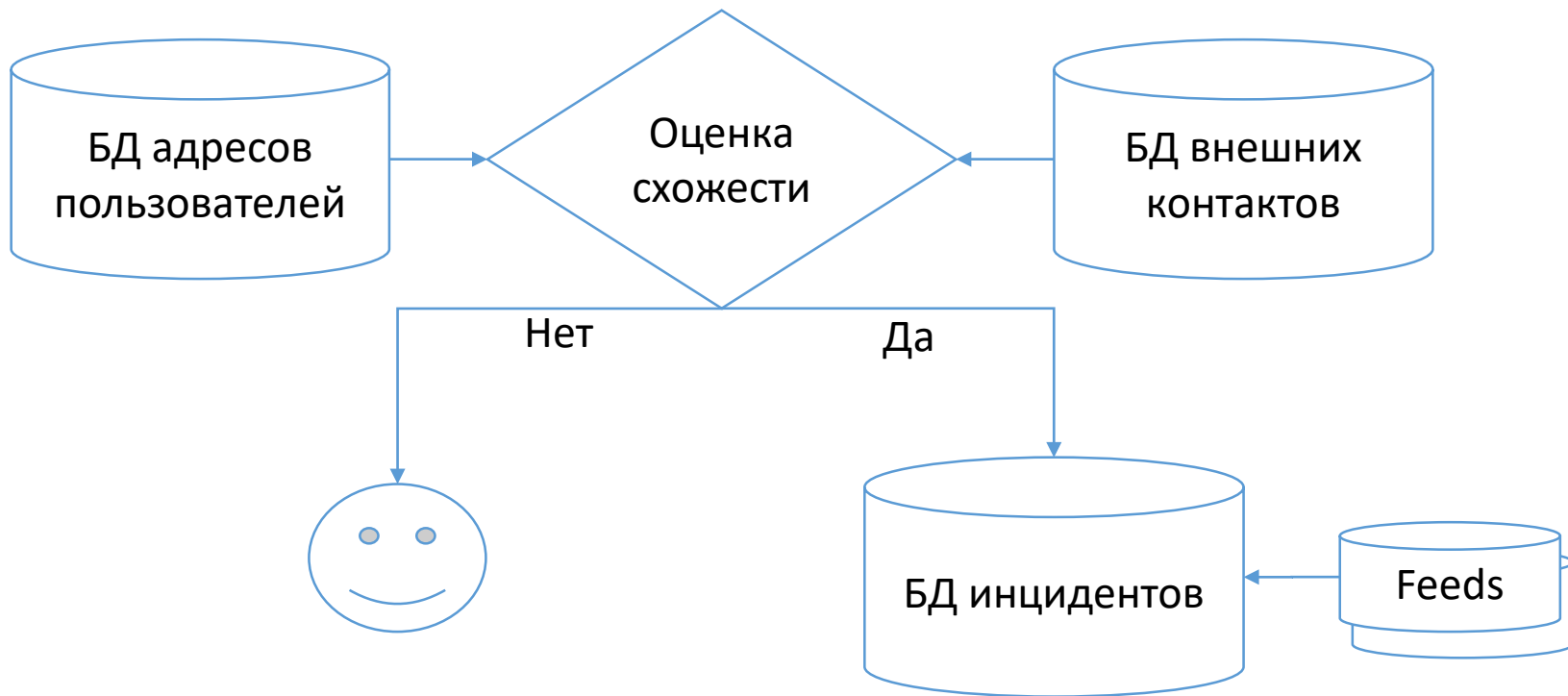
регистрация и bänk.com, bańk.com, baňk.com...

## Решения

- На уровне ICANN – Label Generation Rules
- На уровне Регистратур – запрет кросскрипта и IDN-таблицы
- На уровне Регистраторов – превентивные регистрации
- На уровне разработчиков ПО – средства нотификации пользователей



## Методики



## Оценка схожести доменов или почтовых адресов

строковое расстояние  
(Дамерау — Левенштейна)

|   |   |   |
|---|---|---|
| к | о | т |
| к | о | д |

расстояние = 1

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | к | о | т | и | к |   |
| с | к | о | т | и | н | а |

расстояние = 3

## Нормализация

### Неоднозначность кодирования символов

ñ = U+00F1

ñ = n ~ = U+006E U+007E

niño != niño

## Нормализация

Приведение к каноническому представлению

— Normalization Form D (NFD)

Š (U+1E69) → Š (U+1E61) . (U+0323) → S(U+0073) ˇ (U+0307) . (U+0323)

## Нормализация

Приведение к каноническому представлению

– Normalization Form D (NFD)

– Normalization Form C (NFC)

$\text{NFD} + n_{(U+006E)} \sim_{(U+007E)} \rightarrow \tilde{n}_{(U+00F1)}$

## Нормализация

Приведение к каноническому представлению

- Normalization Form D (NFD)
- Normalization Form C (NFC)
- Normalization Form KD (NFKD)

$\text{™}_{(\text{U}+2122)} \rightarrow \text{T}_{(\text{U}+0054)} \text{M}_{(\text{U}+004\text{D})}$

$\text{◌}_{(\text{U}+\text{FE}37)} \rightarrow \{_{(\text{U}+007\text{B})}$

$\frac{1}{4}_{(\text{U}+00\text{BC})} \rightarrow 1_{(\text{U}+0031)} \text{ / }_{(\text{U}+2044)} 4_{(\text{U}+0034)}$

$\text{Hh}_{(\text{U}+210\text{D})} \rightarrow \text{H}_{(\text{U}+0048)}$

## Нормализация

Приведение к каноническому представлению

- Normalization Form D (NFD)
- Normalization Form C (NFC)
- Normalization Form KD (NFKD)
- Normalization Form KC (NFKC)

NFKD + NFC

## Нормализация

Приведение к каноническому представлению

- Normalization Form D (NFD)
- **Normalization Form C (NFC)**
- Normalization Form KD (NFKD)
- Normalization Form KC (NFKC)

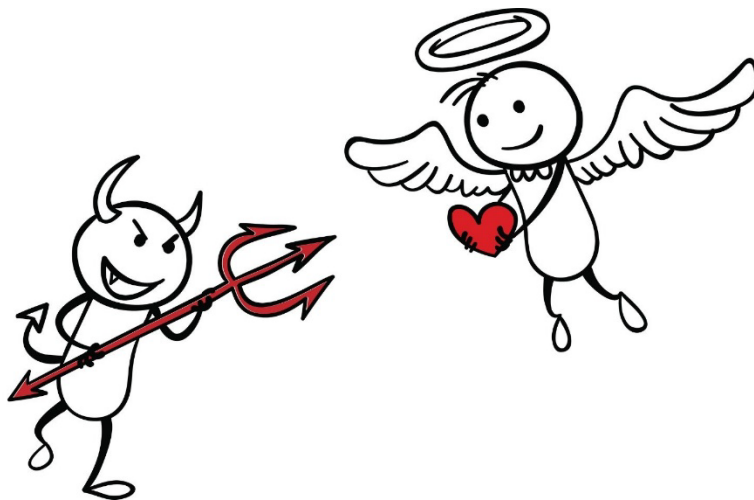
— НОВЫЕ ТЕХНОЛОГИИ = НОВЫЕ РИСКИ

- НОВЫЕ ТЕХНОЛОГИИ = НОВЫЕ РИСКИ
- look-a-like атаки появились раньше IDN и EAI

- НОВЫЕ ТЕХНОЛОГИИ = НОВЫЕ РИСКИ
- look-a-like атаки появились раньше IDN и EAI
- существуют механизмы противодействия
- разрабатываются новые механизмы

- НОВЫЕ ТЕХНОЛОГИИ = НОВЫЕ РИСКИ
- look-a-like атаки появились раньше IDN и EAI
- существуют механизмы противодействия
- разрабатываются новые механизмы
- позволяют существенно снизить риски

## Заклучение



— борьба добра и зла будет всегда...



КООРДИНАЦИОННЫЙ ЦЕНТР  
ДОМЕНОВ .RU/.RF

21

# СПАСИБО ЗА ВНИМАНИЕ!

**Вадим Михайлов**

[mva@cctld.ru](mailto:mva@cctld.ru)

[мва@кц.рф](mailto:mva@кц.рф)