DANMARKS STATISTIK

تشفير المعرفات للمشاريع الخارجية

عمان، 22شباط 2024

أنواع التشفير في الإحصاءات الدنماركية

DANMARKS STATISTIK

- هناك نوعان من التشفير في الإحصاءات الدنماركية :

– إعادة ترقيم عام لعدد الأشخاص إلى person_idفي معظم مجموعات البيانات الإحصائية

- تغيير مقابض رقم الشخص)يحصل الأشخاص أحيانًا على رقم شخص جديد(
- هو رقم)يصل إلى 15خانة(
- عندما يتطلب التبادل مرة أخرى إلى رقم الشخص موعدًا.
- يتم استخدامه لأنه يحدد الأشخاص بشكل أفضل، كما أنه يزيل رقم الشخص من البيانات)رقم الشخص أكثر حساسية(

– تحديد التشفير للمشاريع الخارجية

- اسم مستعار لجميع المتغيرات المحددة)حوالي 150متغيرًا(
- لكل من البيانات من DSTوالبيانات التي قد يجلبها المشروع
- المفتاح محدد لكل مشروع

التشفير للمشاريع الخارجية

- عبارة مرور لكل مشروع .محفوظة في قاعدة بيانات بوابة DDP
- خوارزميتان:
  - القديمة
    - تستخدم في معظم المشاريع
    - قديم .يمكن إرجاعها إلى بطاقتين مثقوبتين في خزانتين مصرفيتين مختلفتين في أوائل عام ." 1990
    - تم تنفيذها في SAS، والتي تولد رمز .SAS
  - الجديدة
    - تم تطويره في 219/2020بسبب الرغبة في عدم الاعتماد على SAS

مثال

DANMARKS STATISTIK

- مفاتيح مختلفة لمشاريع مختلفة

  - Project 1 key: Pass1
  - Project 2 key: Pass2
  - Variable value: 0123456789abcd
  - Pseudonym for variable for project 1: TKlqHWDufwCd8mRJhvTMRA==
  - Pseudonym for variable for project 2: dSeV3K4ryuJj0Mzu0j341w==

الخوارزمية الجديدة

DANMARKS STATISTIK

- تم تطويره في مجال تكنولوجيا المعلومات بالتعاون مع دوائر البحوث
- لا حاجة إلى استخدام SAS
  - عدم استخدام SAS كأداة وعدم استخدام ملفات SAS
    - يجب أن تكون قادرًا على التشغيل على الآلات، حيث لا يوجد ترخيص SAS
    - وهناك أيضا حاجة إلى أن تكون قادرة على التعامل مع أشكال البيانات الجديدة، مثل بيانات الجينوم، مثل (PLINK compact
- السمات والخصائص
  - يجب أن تكون طويلة الأمد)لا تعتمد على الأدوات والخوارزميات التي قد تختفي(
  - لا يحتاج إلى الترخيص، حتى يمكن تشغيله في أي مكان
  - يجب أن تكون الخوارزمية والتنفيذ جاهزين للاستعانة بمصادر مفتوحة)لا تعتمد على الأمان عن طريق الغموض.(

الخوارزمية الجديدة

DANMARKS STATISTIK

- السمات والخصائص
  - استخدم عمليات التشفير المقبولة للثقة
  - قابلاً للتنفيذ بلغات برمجة مختلفة)على سبيل المثال python، c #، java)
  - يجب قبول جميع أنواع المدخلات)نص، ثنائي (وتقديم مخرجات نصية)نص، base64مشفر(
  - الحصول على أداء لائق.
  - تكون قابلة للتطبيق في عدة سياقات، على سبيل المثال مستقل لملفات csvأو كإجراء مخزن في قاعدة بيانات
  - قابل للتكرار
    - القيمة المستعارة بنفس المفتاح تسفر دائمًا عن نفس النتيجة
  - تم تقييم الخوارزمية من قبل مدقق حسابات مستقل pwc)، (2021

جزء من التنفيذ

DANMARKS STATISTIK

```
        _key = _sha_key
    elif AES_KEY_SIZE == 128:
        _key = bytes([_a ^ _b for _a, _b in zip(_sha_key[:AES.block_size], _sha_key[-AES.block_size:])])
    else:
        _key = None
    #
    self.cipher = AES.new(_key, AES.MODE_ECB)
    # Don't leave unneccessary stuff in memory
    password, sha_key, _sha_key, _key, = None, None, None, None

def encrypt(self, s: bytes):
    """The encrypted bitstream is returned base64 encoded"""
    return b64encode(self.cipher.encrypt(pad(s, AES.block_size)))

def decrypt(self, s: bytes):
    """The plaintext is expected to be base64 encoded"""
    return unpad(self.cipher.decrypt(b64decode(s)), AES.block_size)
```

المخرجات

DANMARKS STATISTIK

```
0000000000;ffJbtH2t91rMaocb7fcUdyuw==
0000000001;ffjGro1r/QD8utL19CWNkByw==
0000000002;ff74EuplMs6z3itxV55MOrxA==
0000000003;ffWBzj0y3c9CetQVKwLtTR1A==
0000000004;ffd6nEzIxEfNeYeivA4yKmLA==
0000000005;ffCnIoWnvBEQ1NpRCvd+Xfyg==
0000000006;ff43o9IWudREKVCDFY58CZsA==
0000000007;ffP+G1i+l6yuXwgoanKqc6pw==
0000000008;ffmpUkVZn8EDblLz10UK8t6Q==
0000000009;ffiSE4NKv7lxGfCf4b4DNOaQ==
```