# INTRODUCTION À LA SÉCURITÉ INFORMATIQUE RÉVISIONS

*Marc-Olivier Killijian*
*UQAM, CNRS*
*INF4471 A21*

# K-ANONYMAT

# K-ANONYMAT - SWEENEY

- **Garantir** que dans chaque groupe de l'ensemble assaini, chaque individu est identique à au moins *k-1* autres
  - Généralisation et suppressions
  - Ex: données médicales

| | Non-Sensitive | | | Sensitive |
|---|---|---|---|---|
| | Zip Code | Age | Nationality | Condition |
| 1 | 13053 | 28 | Russian | Heart Disease |
| 2 | 13068 | 29 | American | Heart Disease |
| 3 | 13068 | 21 | Japanese | Viral Infection |
| 4 | 13053 | 23 | American | Viral Infection |
| 5 | 14853 | 50 | Indian | Cancer |
| 6 | 14853 | 55 | Russian | Heart Disease |
| 7 | 14850 | 47 | American | Viral Infection |
| 8 | 14850 | 49 | American | Viral Infection |
| 9 | 13053 | 31 | American | Cancer |
| 10 | 13053 | 37 | Indian | Cancer |
| 11 | 13068 | 36 | Japanese | Cancer |
| 12 | 13068 | 35 | American | Cancer |

**Figure 1. Inpatient Microdata**

| | Non-Sensitive | | | Sensitive |
|---|---|---|---|---|
| | Zip Code | Age | Nationality | Condition |
| 1 | 130** | < 30 | * | Heart Disease |
| 2 | 130** | < 30 | * | Heart Disease |
| 3 | 130** | < 30 | * | Viral Infection |
| 4 | 130** | < 30 | * | Viral Infection |
| 5 | 1485* | ≥ 40 | * | Cancer |
| 6 | 1485* | ≥ 40 | * | Heart Disease |
| 7 | 1485* | ≥ 40 | * | Viral Infection |
| 8 | 1485* | ≥ 40 | * | Viral Infection |
| 9 | 130** | 3* | * | Cancer |
| 10 | 130** | 3* | * | Cancer |
| 11 | 130** | 3* | * | Cancer |
| 12 | 130** | 3* | * | Cancer |

**Figure 2. 4-anonymous Inpatient Microdata**

# ATTAQUE PAR COMPOSITION

- **Question:** le patron d'Alice sait quelle a 28 ans, qu'elle a pour code postal 13012 et qu'elle a visité deux hôpitaux. Que peut-il inférer ?

| | Non-Sensitive | | | Sensitive |
|---|---|---|---|---|
| | Zip code | Age | Nationality | Condition |
| 1 | 130** | <30 | * | AIDS |
| 2 | 130** | <30 | * | Heart Disease |
| 3 | 130** | <30 | * | Viral Infection |
| 4 | 130** | <30 | * | Viral Infection |
| 5 | 130** | ≥40 | * | Cancer |
| 6 | 130** | ≥40 | * | Heart Disease |
| 7 | 130** | ≥40 | * | Viral Infection |
| 8 | 130** | ≥40 | * | Viral Infection |
| 9 | 130** | 3* | * | Cancer |
| 10 | 130** | 3* | * | Cancer |
| 11 | 130** | 3* | * | Cancer |
| 12 | 130** | 3* | * | Cancer |

(a)

| | Non-Sensitive | | | Sensitive |
|---|---|---|---|---|
| | Zip code | Age | Nationality | Condition |
| 1 | 130** | <35 | * | AIDS |
| 2 | 130** | <35 | * | Tuberculosis |
| 3 | 130** | <35 | * | Flu |
| 4 | 130** | <35 | * | Tuberculosis |
| 5 | 130** | <35 | * | Cancer |
| 6 | 130** | <35 | * | Cancer |
| 7 | 130** | ≥35 | * | Cancer |
| 8 | 130** | ≥35 | * | Cancer |
| 9 | 130** | ≥35 | * | Cancer |
| 10 | 130** | ≥35 | * | Tuberculosis |
| 11 | 130** | ≥35 | * | Viral Infection |
| 12 | 130** | ≥35 | * | Viral Infection |

(b)

# AUTRES MÉTRIQUES DE PROTECTION

- Le k-anonymat n'est pas forcément suffisant
- **l-diversité**: maintient de la diversité dans chaque groupe par rapport aux valeurs possibles des attributs sensibles
  - instantiation possible basée sur la mesure de **l'entropie**
  - protection contre les attaques basées sur l'homogénéité
- **t-proximité**: la distribution des attributs dans chaque groupe doit être proche de celle globale
  - *t* est un seuil à ne pas dépasser qui représente la proximité entre distributions

# QUESTION

- Rendre 5 anonyme cette table

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 47677 | 29 | 3K | gastric ulcer |
| 2 | 47602 | 22 | 4K | gastritis |
| 3 | 47678 | 27 | 5K | stomach cancer |
| 4 | 47905 | 43 | 6K | gastritis |
| 5 | 47909 | 52 | 11K | flu |
| 6 | 47906 | 47 | 8K | bronchitis |
| 7 | 47605 | 30 | 7K | bronchitis |
| 8 | 47673 | 36 | 9K | pneumonia |
| 9 | 47607 | 32 | 10K | stomach cancer |

**Table 3. Original Salary/Disease Table**

Source

# QUESTION

- Rendre 5 anonyme cette table en maximisant l'utilité

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 47677 | 29 | 3K | gastric ulcer |
| 2 | 47602 | 22 | 4K | gastritis |
| 3 | 47678 | 27 | 5K | stomach cancer |
| 4 | 47905 | 43 | 6K | gastritis |
| 5 | 47909 | 52 | 11K | flu |
| 6 | 47906 | 47 | 8K | bronchitis |
| 7 | 47605 | 30 | 7K | bronchitis |
| 8 | 47673 | 36 | 9K | pneumonia |
| 9 | 47607 | 32 | 10K | stomach cancer |

**Table 3. Original Salary/Disease Table**

# QUESTION

- Quelles attaques restent possibles ?

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 47677 | 29 | 3K | gastric ulcer |
| 2 | 47602 | 22 | 4K | gastritis |
| 3 | 47678 | 27 | 5K | stomach cancer |
| 4 | 47905 | 43 | 6K | gastritis |
| 5 | 47909 | 52 | 11K | flu |
| 6 | 47906 | 47 | 8K | bronchitis |
| 7 | 47605 | 30 | 7K | bronchitis |
| 8 | 47673 | 36 | 9K | pneumonia |
| 9 | 47607 | 32 | 10K | stomach cancer |

**Table 3. Original Salary/Disease Table**

# QUESTION

- Quelles attaques restent possibles ?
  - Homogénéité
  - Connaissance auxiliaire

| | ZIP Code | Age | Disease |
|---|---|---|---|
| 1 | 47677 | 29 | Heart Disease |
| 2 | 47602 | 22 | Heart Disease |
| 3 | 47678 | 27 | Heart Disease |
| 4 | 47905 | 43 | Flu |
| 5 | 47909 | 52 | Heart Disease |
| 6 | 47906 | 47 | Cancer |
| 7 | 47605 | 30 | Heart Disease |
| 8 | 47673 | 36 | Cancer |
| 9 | 47607 | 32 | Cancer |

**Table 1. Original Patients Table**

| | ZIP Code | Age | Disease |
|---|---|---|---|
| 1 | 476** | 2* | Heart Disease |
| 2 | 476** | 2* | Heart Disease |
| 3 | 476** | 2* | Heart Disease |
| 4 | 4790* | $\geq 40$ | Flu |
| 5 | 4790* | $\geq 40$ | Heart Disease |
| 6 | 4790* | $\geq 40$ | Cancer |
| 7 | 476** | 3* | Heart Disease |
| 8 | 476** | 3* | Cancer |
| 9 | 476** | 3* | Cancer |

**Table 2. A 3-Anonymous Version of Table 1**

# QUESTION

- Rendre cette table 3-diverse

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 47677 | 29 | 3K | gastric ulcer |
| 2 | 47602 | 22 | 4K | gastritis |
| 3 | 47678 | 27 | 5K | stomach cancer |
| 4 | 47905 | 43 | 6K | gastritis |
| 5 | 47909 | 52 | 11K | flu |
| 6 | 47906 | 47 | 8K | bronchitis |
| 7 | 47605 | 30 | 7K | bronchitis |
| 8 | 47673 | 36 | 9K | pneumonia |
| 9 | 47607 | 32 | 10K | stomach cancer |

**Table 3. Original Salary/Disease Table**

# QUESTION

- Rendre cette table 3-diverse

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 476** | 2* | 3K | gastric ulcer |
| 2 | 476** | 2* | 4K | gastritis |
| 3 | 476** | 2* | 5K | stomach cancer |
| 4 | 4790* | $\geq 40$ | 6K | gastritis |
| 5 | 4790* | $\geq 40$ | 11K | flu |
| 6 | 4790* | $\geq 40$ | 8K | bronchitis |
| 7 | 476** | 3* | 7K | bronchitis |
| 8 | 476** | 3* | 9K | pneumonia |
| 9 | 476** | 3* | 10K | stomach cancer |

**Table 4. A 3-diverse version of Table 3**

# QUESTION

- Il reste des attaques possibles
  - Liées à la proximité sémantique
  - Si je sais que Bob a la vingtaine (ou je connais son CP) alors je sais qu'il gagne peu et qu'il a des problèmes d'estomac

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 476** | 2* | 3K | gastric ulcer |
| 2 | 476** | 2* | 4K | gastritis |
| 3 | 476** | 2* | 5K | stomach cancer |
| 4 | 4790* | $\geq 40$ | 6K | gastritis |
| 5 | 4790* | $\geq 40$ | 11K | flu |
| 6 | 4790* | $\geq 40$ | 8K | bronchitis |
| 7 | 476** | 3* | 7K | bronchitis |
| 8 | 476** | 3* | 9K | pneumonia |
| 9 | 476** | 3* | 10K | stomach cancer |

**Table 4. A 3-diverse version of Table 3**

# QUESTION

- Il y a la t-proximité qui mesure le gain d'information dans une classe
- La différence de distance entre les valeurs d'un attribut dans une classe et toutes classes comprises doit être de t maximum

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 476** | 2* | 3K | gastric ulcer |
| 2 | 476** | 2* | 4K | gastritis |
| 3 | 476** | 2* | 5K | stomach cancer |
| 4 | 4790* | $\geq 40$ | 6K | gastritis |
| 5 | 4790* | $\geq 40$ | 11K | flu |
| 6 | 4790* | $\geq 40$ | 8K | bronchitis |
| 7 | 476** | 3* | 7K | bronchitis |
| 8 | 476** | 3* | 9K | pneumonia |
| 9 | 476** | 3* | 10K | stomach cancer |

**Table 4. A 3-diverse version of Table 3**

# QUESTION

- Il y a la t-proximité qui mesure le gain d'information dans une classe
- La différence de distance entre les valeurs d'un attribut dans une classe et toutes classes comprises doit être de t maximum

| | ZIP Code | Age | Salary | Disease |
|---|---|---|---|---|
| 1 | 4767* | $\leq 40$ | 3K | gastric ulcer |
| 3 | 4767* | $\leq 40$ | 5K | stomach cancer |
| 8 | 4767* | $\leq 40$ | 9K | pneumonia |
| 4 | 4790* | $\geq 40$ | 6K | gastritis |
| 5 | 4790* | $\geq 40$ | 11K | flu |
| 6 | 4790* | $\geq 40$ | 8K | bronchitis |
| 2 | 4760* | $\leq 40$ | 4K | gastritis |
| 7 | 4760* | $\leq 40$ | 7K | bronchitis |
| 9 | 4760* | $\leq 40$ | 10K | stomach cancer |

**Table 5. Table that has 0.167-closeness w.r.t. Salary and 0.278-closeness w.r.t. Disease**

# ARITHMÉTIQUE MODULAIRE

# 101

- $25 \equiv 1 \mod 3$ : quand on divise 25 par 3, le reste est 1
- $A \equiv R \mod m$ : quand on divise A par R, le reste est m
- $A \equiv 0 \mod m$ : si $m$ est diviseur de $A$
- *Méthode empirique :* $A \equiv R \mod m$
  - $A/m=x.y \Rightarrow R=A-xm \Rightarrow A \equiv (A - xm) \mod m$

# 101

- $25 \equiv 1 \mod 3$
- Mais aussi $25 \equiv 4 \mod 3$
- Et aussi $25 \equiv 300000001 \mod 3$

# QUESTIONS

- Calculer 123456789 mod 9876

# QUESTIONS

- Calculer 123456789 mod 9876
- 123456789/9876=12500.6874...
- 123456789-9876*12500 = 6789
- 123456789 ≡ 6789 mod 9876 palindromesque !

# RSA

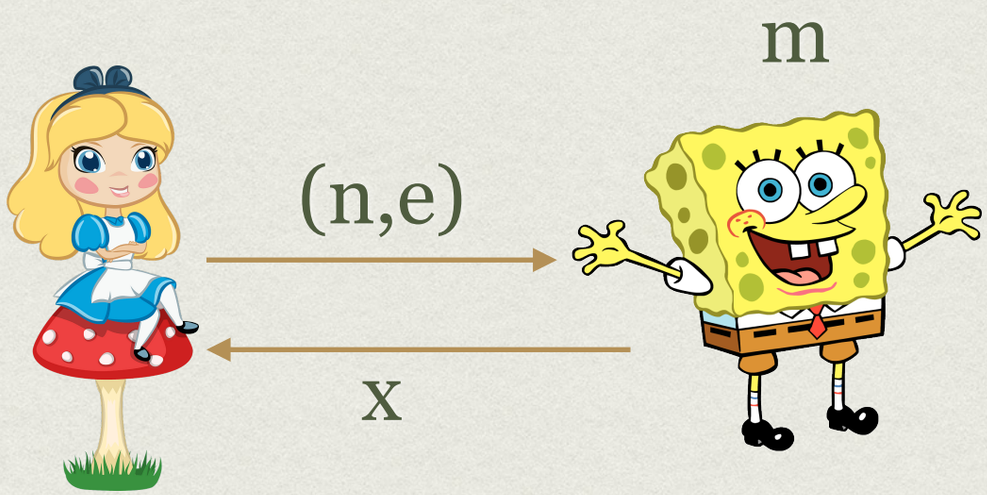
# RSA.1 PRÉPARATION DES CLÉS

Clé publique = (n,e)

Clé privée = d

1. Choix de deux premiers distincts $p$ et $q$
2. Calcul de $n = pq$
3. Calcul de $\phi(n) = (p-1)(q-1)$
4. Choix d'un exposant e tel que $pgcd(e, \phi(n)) = 1$
5. Calcul de $d$ l'inverse de e modulo $\phi(n)$ par Euclide étendu $d.e \equiv 1 \bmod \phi(n)$

# RSA.2 CHIFFREMENT

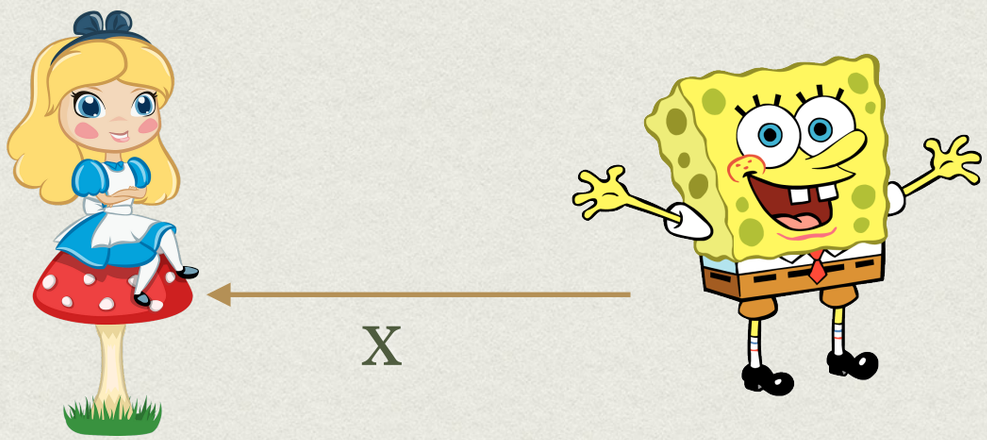



1. Bob veut envoyer un message à Alice
   Récupère la clé publique (n,e)
   Encodage d'un ou plusieurs messages $m_i < n$

2. Bob chiffre m par exponentiation rapide
   $x \equiv m^e \bmod n$

3. Il transmet le chiffré x à Alice

# RSA.3 DÉCHIFFREMENT




1. Alice reçoit x de Bob

2. Alice déchiffre x par exponentiation rapide
   $m \equiv x^d \bmod n$

# QUESTION

- On choisit p=7, q=9
- Est-ce que e=3 est un bon exposant ?
  - Calculer n, $\phi(n)$, $pgcd(e, \phi(n)) = 1$

# QUESTION

- On choisit p=7, q=9
- Est-ce que e=3 est un bon exposant ?
  - Calculer n, $\phi(n)$, 3 est-il premier avec $\phi(n)$

# QUESTION

- On choisit p=7, q=9
- Est-ce que e=4 est un bon exposant ?
  - Calculer n, $\phi(n), pgcd(e, \phi(n)) = 1$

# QUESTION

- On choisit p=7, q=9
- Est-ce que e=5 est un bon exposant ?
  - Calculer n, $\phi(n), pgcd(e, \phi(n)) = 1$

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- Calculer $d = e^{-1} \bmod \phi(n)$
  - PGCD, Bézout ...

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- Calculer $d = e^{-1} \bmod \phi(n)$
  - PGCD, Bézout ...
  - d=-19 ....

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- Calculer $d = e^{-1} \bmod \phi(n)$
  - PGCD, Bézout ...
  - $d \equiv -19 \equiv 29 \bmod 48$

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- d=29
- Chiffrer le message m=67

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- d=29
- Non, 67 >n il faudrait encoder m en 2 messages <n
- Chiffrer m=52

# QUESTION

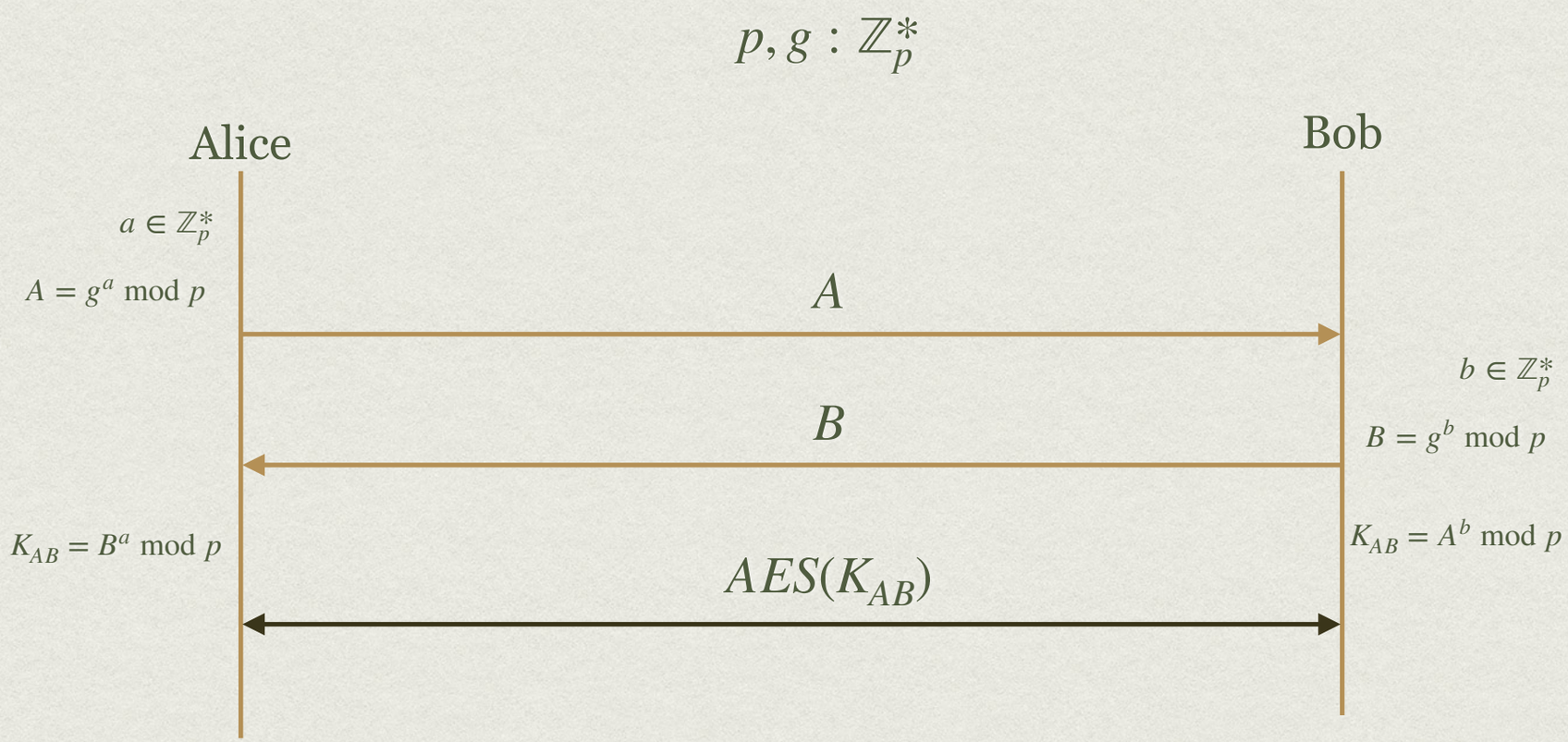
- On choisit p=7, q=9
- On choisit e=5
- d=29
- Non, 67 >n il faudrait encoder m en 2 messages <n
- Chiffrer m=52 (x=55)
- Déchiffrer x=3

# QUESTION

- On choisit p=7, q=9
- On choisit e=5
- d=29
- Non, 67 >n il faudrait encoder m en 2 messages <n
- Chiffrer m=52 (x=55)
- Déchiffrer x=3 (m=54)

# DIFFIE-HELLMAN

# ECHANGE DE CLÉ DIFFIE-HELLMAN




Preuve: $K_{AB} = B^a \bmod p = (\alpha^b)^a \bmod p = (\alpha^a)^b \bmod p = A^b \bmod p = K_{AB}$

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 13$
- Vous tirez b=14, quel B envoyez vous à Alice ?

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 13$
- Vous tirez b=14, vous envoyez $B = 2^{14} \bmod 19$=6

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 13$
- Vous tirez b=14, vous envoyez $B = 2^{14} \bmod 19$=6
- Quelle est la clé symétrique ?

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 2$
- Vous tirez b=14, vous envoyez $B = 2^{14} \bmod 19$=6
- Quelle est la clé symétrique ?
  $K = B^a = A^b \bmod 19 = 2^{14} \bmod 19$=6

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 2$
- Vous tirez b=14, vous envoyez $B = 2^{14} \bmod 19$=6
- Quelle est la clé symétrique ?
  $K = B^a = A^b \bmod 19 = 2^{14} \bmod 19$=6
- Pouvez vous deviner le nombre secret d'Alice ?

# QUESTION

- On choisit p=19 et g=2
- Vous êtes Bob, Alice vous envoie $A = g^a \bmod p = 2$
- Vous tirez b=14, vous envoyez $B = 2^{14} \bmod 19$=6
- Quelle est la clé symétrique ?
  $K = B^a = A^b \bmod 19 = 2^{14} \bmod 19$=6
- Avec de tels petits nombres, oui, car vous pouvez énumérer :
  $A = g^a \bmod p = 2^a \bmod 19 = 2 \ldots 2^a = 2$ donc ...

# ET LE RESTE ?

- Questions de compréhension du cours
- Environ 1 question par section du cours
- (9h30-12h30)=3h
- Arrivez à l’heure
- Ayez des brouillons pour les calculs

« Bon courage et bonne fêtes de fin d’année »

*-Marc-Olivier K.*