AUTOR
António Machiavelo
Univ. do Porto
ajmachia@fc.up.pt

SEGURANÇA PRÉ E PÓS-QUÂNTICA

O uso de códigos secretos, ou cifras, para transmitir informações confidenciais é multimilenário. O advento da Internet multiplicou os usos da criptografia, e hoje prepara-se já o futuro da segurança informática num potencial mundo povoado por uma nova geração de computadores radicalmente diferentes dos atuais – os computadores quânticos.

A *criptologia*, o estudo de cifras, isto é, de métodos seguros de transmissão de informação confidencial por forma a que esta seja ilegível por terceiros, engloba duas atividades algo opostas mas inseparáveis: a *criptografia*, a criação de cifras, e a *criptanálise*, a procura sistemática de fragilidades nessas mesmas cifras. Para desenhar uma cifra segura é necessário considerar a sua criptanálise e para fazer criptanálise é essencial conhecer os sistemas criptográficos e as ideias subjacentes à construção de cada um deles.

A criptografia é bem antiga, remontando provavelmente às primeiras civilizações humanas [9, Cap.2]. Mas o estudo sistemático da criptologia, e em particular da criptanálise, parece ter sido iniciado no seio do império árabe, provavelmente com importantes antecedentes persas [9, pp. 93-99].

É também antigo o envolvimento de matemáticos na criptologia. Por exemplo, François Viète (1540-1603), famoso pelo seu tratamento da teoria das equações, criptanalisou mensagens secretas espanholas ao serviço do rei Henrique IV de França [9, pp. 116-118]. Mais recentemente, Alan Turing (1912-1954), um dos fundadores da ciência de computadores, desempenhou um papel fundamental na criptanálise da cifra *Enigma*, um sistema criptográfico usado pelas tropas alemãs na Segunda Guerra Mundial [10, 12].

Todas as cifras *clássicas*, isto é, anteriores a 1976, são *simétricas*, ou seja, sabendo a chave usada para cifrar, sabe-se também como decifrar. Um exemplo simples e bem conhecido consiste em usar uma permutação pré-combinada das letras do alfabeto. Hoje as cifras clássicas têm apenas um interesse mais ou menos lúdico, embora a sua criptanálise possa constituir um bom exercício [11] e há várias ideias essenciais na criptologia clássica que são ainda hoje pertinentes.

Em 1976, Whitfield Diffie e Martin Hellman operaram uma mudança de paradigma em criptografia. Propuseram um método engenhoso para duas entidades selecionarem uma chave secreta, de um modo seguro, através de um canal não seguro – método hoje conhecido por *protocolo de Diffie-Hellman* –, e introduziram o conceito de *criptografia de chave pública*, explicando como poderiam ser elaborados sistemas criptográficos em que a chave usada para decifrar é, de um ponto de vista computacional, praticamente impossível de obter conhecendo a chave usada para cifrar, que pode pois ser pública [3]. Apesar de não apresentar nenhuma proposta concreta deste novo tipo de criptografia, o artigo estimulou a procura de um tal sistema e, em 1978, Ronald Rivest, Adi Shamir e Leonard Adleman publicaram o primeiro sistema prático de chave pública [13], que pode ser usado para transmitir informação confidencial e também como um esquema de assinatura digital, que ficou conhecido pela sigla formada pelas iniciais dos apelidos dos seus inventores: RSA.

A criptografia de chave pública veio alargar o uso da criptografia, que, se antes se limitava essencialmente a aplicações militares e a proteger segredos diplomáticos e industriais, hoje é ubíqua, protegendo transações feitas via

Internet, integridade e autenticação de documentos e entidades, assim como os mais diversos dados pessoais dos utilizadores.

O protocolo de Diffie-Hellman e a cifra RSA usam resultados de Teoria dos Números descobertos e explorados nos séculos XVII e XVIII. Dados[1] $n, a \in \mathbb{N}$, vamos designar aqui por $r_n(a)$ o resto da divisão de a por n. Por exemplo, $r_3(14) = 2$. No século XVII, Pierre de Fermat descobriu o seguinte resultado notável: se p é um número primo e $a \in \mathbb{N}$ é um número não divisível por p, então $r_p(a^{p-1}) = 1$ (sempre!). Por razões que não cabe aqui detalhar, Fermat não deixou escrita nenhuma demonstração deste, como de muitos outros resultados (na época não existiam ainda revistas científicas). Coube a Leonhard Euler, que viveu no século XVIII, fornecer demonstrações deste resultado. É claro que bastava uma para garantir a sua veracidade, mas diferentes demonstrações podem fornecer perspetivas extras que permitem ir mais fundo. Foi precisamente assim que Euler acabou por descobrir uma notável generalização do resultado de Fermat: dado $n \in \mathbb{N}$, seja $\varphi(n) = \#\{k \in \mathbb{N} : \mathrm{mdc}(k, n) = 1\}$; então $r_n(a^{\varphi(n)}) = 1$ sempre(!) que $\mathrm{mdc}(a, n) = 1$. Euler descobriu ainda que para cada número primo p existem números g tais que os números $r_p(g^i)$, para $i = 1, 2, \ldots, p-1$, dão exatamente todos os números de $\{1, 2, \ldots, p-1\}$ (por uma outra ordem). Estes números são chamados *raízes primitivas módulo p*. Acontece que há algoritmos eficazes e rápidos (em certo sentido) que permitem calcular $s = r_p(g^i)$, dados p, g, i, mas não se conhecem algoritmos "eficientes" para calcular i dados p, g, s – questão esta (a do cálculo de i à custa de p, g, s) que é conhecida pelo nome *problema do logaritmo discreto*, pois i é uma espécie de logaritmo (numa versão discreta, pois há apenas um número finito de possibilidades) de s módulo p, relativo à base g.

O protocolo de Diffie-Hellman pode agora ser descrito do seguinte modo. Duas entidades A e B que queiram escolher, através de um canal não seguro, uma chave secreta para ser usada numa cifra simétrica, começam por escolher um primo p e uma sua raiz primitiva g, sendo p suficientemente grande para a resolução do respetivo problema do logaritmo discreto ser um sério obstáculo a potenciais espiões. De seguida, A escolhe (de modo aleatório) um número $a \in \{2, 3, \ldots, p-3\}$ e calcula $u = r_p(g^a)$, enquanto B escolhe b nas mesmas condições e calcula $v = r_p(g^b)$. Depois, através do canal (que pode estar sob escuta), A envia u a B, e este envia v a A. Finalmente, A calcula $r_p(v^a)$ e B calcula $r_p(u^b)$. Ambos acabam por obter o mesmo número, k, pois $r_p(v^a) = r_p((g^b)^a) = r_p(g^{ba}) = r_p(g^{ab}) = r_p((g^a)^b) = r_p(u^b)$. O número k pode, pois, ser usado como chave secreta para comunicações futuras, uma vez que quem escutou toda a transmissão, não conhecendo nem a, nem b (a menos que consiga calcular logaritmos discretos), não consegue obter k.

Por seu lado, a cifra RSA baseia-se na observação de que, se $n = pq$, sendo p e q dois primos distintos, e se c for escolhido de tal modo que $\mathrm{mdc}(c, \varphi(n)) = 1$, então existe um número d tal que $r_{\varphi(n)}(cd) = 1$, obtendo-se duas funções, $x \mapsto r_n(x^c)$ e $x \mapsto r_n(x^d)$, do conjunto $\{1, 2, \ldots, n-1\}$ nele próprio, que são inversas uma da outra – facto que resulta do teorema de Euler atrás mencionado. A primeira delas pode, pois, ser usada para cifrar mensagens por todos os que conheçam n e c, a chave dita *pública*, e a outra para decifrar mensagens por quem conheça a *chave privada d*. Já agora, d é calculado a partir de c e de $\varphi(n)$, usando um método com mais de 2000 anos, conhecido pelo nome de *algoritmo estendido de Euclides*.

Enquanto o protocolo de Diffie-Hellman baseia a sua segurança na dificuldade de calcular logaritmos discretos, a segurança do sistema RSA assenta na dificuldade de fatorizar números a partir de um certo tamanho, quando comparada com a relativa facilidade em gerar números primos da mesma ordem de grandeza. Por exemplo, enquanto se pode hoje gerar primos com 1000 algarismos em frações de segundo, a fatorização de números como o seguinte:

25 195 908 475 657 893 494 027 183 240
048 398 571 429 282 126 204 032 027 777
137 836 043 662 020 707 595 556 264 018 525
880 784 406 918 290 641 249 515 082 189 298
559 149 176 184 502 808 489 120 072 844 992
687 392 807 287 776 735 971 418 347 270
261 896 375 014 971 824 691 165 077 613 379 859
095 700 097 330 459 748 808 428 401 797 429
100 642 458 691 817 195 118 746 121 515 172
654 632 282 216 869 987 549 182 422 433 637
259 085 141 865 462 043 576 798 423 387 184
774 447 920 739 934 236 584 823 824 281 198
163 815 010 674 810 451 660 377 306 056 201
619 676 256 133 844 143 603 833 904 414 952
634 432 190 114 657 544 454 178 424 020 924
616 515 723 350 778 707 749 817 125 772 467
962 926 386 356 373 289 912 154 831 438 167
899 885 040 445 364 023 527 381 951 378 636
564 391 212 010 397 122 822 120 720 357,

[1] Como é usual em Teoria dos Números, $\mathbb{N}$ designa o conjunto dos números naturais {1, 2, 3, 4, 5, ...}.

está fora do alcance dos nossos melhores algoritmos e de todo o poder computacional existente. Este número, que é conhecido por RSA2048, tem o tamanho de um "*n*" típico usado para assegurar a transação de movimentos financeiros via Internet, por exemplo.

Em 1985, Taher Elgamal mostrou que a ideia em que assenta o protocolo de Diffie e Hellman pode ser adaptada para dar uma cifra e um método de assinatura digital [4]. No mesmo ano, e de um modo independente, Victor S. Miller e Neal Koblitz sugerem o uso de *curvas elípticas* – certas curvas dadas por polinómios cúbicos de duas variáveis que podem ser munidas de uma estrutura de *grupo*, ou seja, de uma operação que a dois pontos da curva associa um outro ponto e que é associativa, tem elemento neutro e cada elemento tem um inverso. A vantagem do uso das curvas elípticas sobre os números primos e raízes primitivas da proposta original de Diffie e Hellman advém do facto de haver, para cada primo, várias curvas elípticas definidas sobre esse primo (i.e., para quem sabe o que isso significa, definida sobre o corpo com esse número primo de elementos).

Porém, em 1994, Peter Shor mostrou que num computador quântico – um computador que tira proveito de certos fenómenos subatómicos – podem implementar-se algoritmos que, em certo sentido, "resolvem" o problema da fatorização e o problema do logaritmo discreto [14]. Ou seja, um computador quântico torna obsoletos os protocolos de Diffie-Hellman e as cifras RSA e Elgamal. Apesar de não haver ainda computadores quânticos merecedores desse nome, sendo as dificuldades tecnológicas da sua construção imensas, há já alguns protótipos relativamente

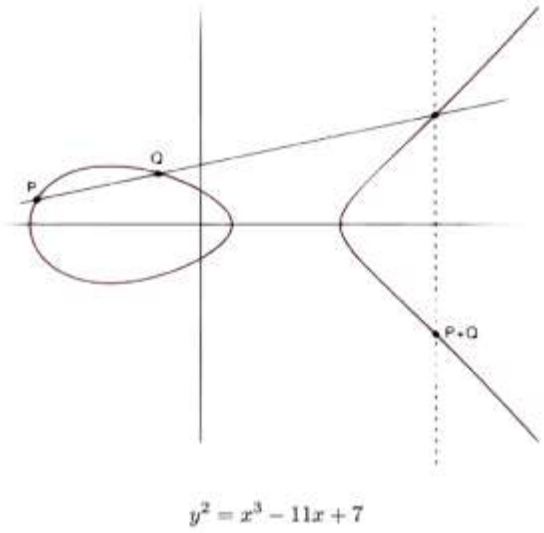

Figura 1. Um exemplo de soma de pontos numa curva elíptica

rudimentares que são suficientes para causar sérias preocupações sobre o futuro da segurança informática.

Em consequência disso, tem havido um enorme interesse no desenvolvimento e no estudo de sistemas criptográficos alternativos que se pensa poderem resistir a ataques quânticos (em criptologia, como na vida, as certezas absolutas são muito escassas). Há várias propostas destes sistemas ditos *pós-quânticos*, usando as mais variadas ferramentas matemáticas [2]: códigos corretores de erros, reticulados geométricos, sistemas de equações quadráticas de várias variáveis. Há também propostas recentes interessantes usando grupos não-abelianos [7, 8], i.e. grupos onde a operação não é comutativa.

Uma outra alternativa que só muito recentemente [1, 15] se começou a explorar com alguma profundidade deriva do trabalho do criptógrafo chinês Renji Tao, e assenta na ideia de que a fatorização de certas composições de *autómatos finitos* – modelos matemáticos abstratos, simples mas bastante profícuos, de computação – poderá ser computacionalmente intratável (em certo sentido).

Na introdução do seu artigo intitulado "Teoremas sobre os Divisores dos Números" [5], L. Euler escreveu:

> " [...] *o conhecimento de uma qualquer verdade vale a pena por si só, mesmo que esta pareça não relacionada com o quotidiano; vimos já que todas as verdades, pelo menos as que conseguimos compreender, estão tão fortemente interligadas entre si que não podemos considerar uma qualquer delas de todo inútil sem alguma imprudência.*
>
> *E portanto, mesmo que uma determinada proposição pareça ser tal que, sendo verdadeira ou falsa, não nos traga absolutamente nenhum benefício, mesmo assim, o método pelo qual se venha a estabelecer a sua verdade ou falsidade, pode, no entanto, ser útil na abertura de caminhos para a descoberta de outras verdades mais úteis. Por esta razão, acredito firmemente que não gastei inutilmente o meu trabalho e o meu esforço na investigação das demonstrações de certas proposições. Por conseguinte, esta teoria de divisores não carece de qualquer uso, mas pode, em algum momento futuro, mostrar uma utilidade que não pode ser desprezada.*
>
> *Estou também especialmente convicto de que o método de cálculo que aqui uso pode, em algum momento, contribuir de um modo não desprezável para investigações mais sérias.*"

É muito interessante observar que aquilo que Euler expõe neste artigo, após estas observações introdutórias, é a sua

segunda demonstração do resultado de Fermat acima mencionado. Foi analisando as descobertas de Fermat com cuidado, e profundidade, que Euler viria a descobrir a generalização que acima ficou descrita (que ele expôs em [6]) e que está na base do RSA, assim como os resultados nos quais assenta o protocolo de Diffie-Hellman. Euler não podia estar mais certo nas suas palavras! E pode-se argumentar que as ideias expostas neste seu artigo, e em outros subsequentes, motivaram ideias em Teoria de Grupos que hoje se usam para tentar resolver o problema de segurança colocado por eventuais computadores quânticos. Este é mais um exemplo da importância da pesquisa dita fundamental, da relação simbiótica entre matemática pura e aplicada, muitas vezes ignorada ou menosprezada pela sociedade em geral e pelas entidades financiadoras em particular.

Espero ter aqui ilustrado, ainda que brevemente, o quanto a matemática é fulcral na criptologia moderna e que, se a única coisa certa acerca do futuro é que ele é incerto, isso não significa que deixemos de tentar antecipá-lo e de nos precavermos da melhor forma possível para o que há de vir. A história do conhecimento, sendo a história que aqui se resumiu um minúsculo exemplo, mostra que tentar até pode dar resultado, e frutos interessantes.

REFERÊNCIAS

[1] Ivone Amorim, *Linear Finite Transducers Towards a Public Key Cryptographic System*, tese de Doutoramento, Faculdade de Ciências da Universidade do Porto, 2016.

[2] Daniel J. Bernstein, Tanja Lange, *Post-Quantum Cryptography – dealing with the fallout of physics success*, Cryptology ePrint Archive, 2017, Report 2017/314, disponível em *https://eprint.iacr.org/2017/314*.

[3] W. Diffie, M. Hellman, *New Directions in Cryptography*, IEEE Transactions on Information Theory, IT-22, nov. 1976, pp. 644- 654.

[4] Taher Elgamal, *A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms*, IEEE Transactions on Information Theory, IT-31, July 1985, pp. 469- 472.

[5] L. Euler, *Theoremata Circa Divisores Numerorum* (E134). Novi Commentarii academiae scientiarum Petropolitanae 1 (1750), pp. 20-48. Reimpressso em *Opera Omnia*: Séries 1, Vol. 2, pp. 63-85. O artigo original, juntamente com uma tradução para o inglês de David Zhao, está disponível em *www.eulerarchive.org*.

[6] L. Euler, *Theoremata Arithmetica Nova Methodo Demonstrata* (E271). Novi Commentarii academiae scientiarum Petropolitanae 8 (1763), pp. 74-104. Reimpressso em *Opera Omnia*: Series 1, Vol. 2, pp. 531-555. O artigo original, juntamente com uma tradução para o alemão de Artur Diener e Alexander Aycock, está disponível em *www.eulerarchive.org*.

[7] Ramón Flores, Delaram Kahrobaei, *Cryptography with Right-Angled Artin Groups*, Theoretical and Applied Informatics, 2016, Vol. 28, N.º3, pp. 8-16.

[8] Jonathan Gryak, Delaram Kahrobaei, *The Status of Polycyclic Group-Based Cryptography: a Survey and Open Problems*, Groups Complexity Cryptology, 2016, Vol. 8, N.º 2, pp. 171-186.

[9] David Kahn, *The Codebreakers: the story of secret writing*, Scribner, 1967.

[10] A. Machiavelo, "ENIGMA: uma história que devia ser mais conhecida", *Gazeta de Matemática*, 2004, N.º 147, pp. 14-15.

[11] A. Machiavelo e Rogério Reis, *Automated Ciphertext-Only Cryptanalysis of the Bifid Cipher*, Cryptologia, 2007, Vol. 31, N.º 2, pp. 112-124.

[12] A. Machiavelo e Rogério Reis,"Turing e a Enigma", *Boletim da SPM*, 2012, Vol. 67, pp. 97-120.

[13] R. Rivest, A. Shamir, L. Adleman, *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM, 1978, Vol. 21, pp. 120-126.

[14] Peter W. Shor, *Algorithms for Quantum Computation: Discrete Logarithms and Factoring*, in: Proceedings, 35th Annual Symposium on Foundations of Computer Science, Santa Fé, NM, November 20-22, 1994, IEEE Computer Society Press, pp. 124-134.

[15] Joana Vieira, *Finite Transducers in Public Key Cryptography*, tese de Mestrado, Faculdade de Ciências da Universidade do Porto, 2017.

Coordenação do espaço PT-MATHS-IN:
Paula Amaral, Universidade Nova de Lisboa, *pt-maths-in@ spm.pt.*