De Bruijn Sequences using properties of finite fields

Cryptography Homework-Andreea Moldovan [322/2]

ChristmasTime 2021

1 A little 'bit' about De Bruijn sequences

As we already know, a de Bruijn sequence of order n is a cyclic sequence of length 2^n,where each substring of length n is a unique binary string.
For example: the sequence 0000100110101111 (of length 16) is a de Bruijn sequence for n = 4. The 16 unique substrings of length 4 when considered cyclicly are:

$$0000, 0001, 0010, 0100, 1001, 0011, 0110, 1101, 1010, 0101, 1011, 0111, 1111, 1110, 1100, 1000$$

As illustrated in this example, a de Bruijn sequence of order n induces a very specific type of cyclic order of the length n binary strings: the length n-1 suffix of a given binary string is the same as the length n-1 prefix of the next string in the ordering.

Another example that contains each length-2 sequence from {0,1,2} exactly once.
We have the following result:

$$0011221020$$

from which we extract the sequence:

$$00, 01, 11, 12, 22, 21, 10, 02, 20$$

and what we get after 'shortening' it:

$$001122102$$

(this is the De Bruijn sequence(circular))

The problem that arises ,though, is: *Do such sequences exist for an arbitrary length over an arbitrary alphabet? And if so: how do we construct them?*
The great news is that if we can find an algorithm to construct such sequences, we answer both questions at the same time!There are many approaches. For example using:

- directed graphs (as I presented last time :)))
- finite fields

In the next section we will study the second.

2 The Construction via finite fields

Suppose we have to find a de Bruijn sequence for the set {0,1,2}, with the given length:2. We consider the polynomial *q(x)=*x^2*+x+2* over Z_3)
Starting with the polynomial $h_0(x)=x$ and $h_{n+1}(x)$ is obtained as it follows: h_n(x)*x/(mod)q(x) (we multiply by x and reduce modulo q(x))
Doing this, we generate the following polynomial sequence(over Z_3)

$$x, 2x + 1, 2x + 2, 2, 2x, x + 2, x + 1, 1$$

(at this point it starts repeating)
We generate a sequence by choosing a degree and taking the corresponding coefficient in each
in our case, we can choose x, so we have

12202110

which is almost a de Bruijn sequence(it misses the 00 but we can add one 0)
we obtain:

012202110

With leading 0:

001220211

2.0.1 Why/How does this work?

It is widely known that for any prime p the ring of integers modulo p, denoted by Z_p, is a field(since it has no zero divisors)
Let
q(x)=$x^l - a_{l-1}x^{l-1} - a_{l-2}x^{l-2} - ... - a_1x - a_0$
be a polynomial of degree l in Z_3 that is irreducible over Z_p, and consider the quotient ring Z_p[x]/q(x)
We know, because q(x) is irreducible, Z_p[x]/q(x) is a field with p^l elements. This is called the *Galois field of order* p^l, often denoted by GF(p^l)

The addition in this field amounts to the addition in Z_p[x]. In multiplying though, we need to reduce modulo q(x).
We do this using the relation:
x^l=$a_{l-1}x^{l-1} + a_{l-2}x^{l-2} + ... + a_1x + a_0$
Observation:
We can view GF(p^l) as a vector space of dimension l over Z_p(Z_p[x] is a vector space over Z_p and this field is a quotinent vector space)
So, the field GF(p^l) is the unique field with p^l elements and it contains all the roots of q(x).
The GF(p^l)* (the set of nonzero elements of the field) forms a cyclic group under multiplication. Suppose we choose q(x) in such a way that one of it roots α is a generator of the cyclic group (so: GF(p^l)*={α,α^2,...,α^{p^l-1}} (this is called a primitive root)
This is exactly what we did in our example, where p=3, l=2, q(x)=$x^2 + x + 2$. Since all elements of GF(3^2) can be viewed as polynomials of degree at most 1, we choose the polynomial $\alpha = x$ as the primitive root. We see that powers of x generate GF(3^2)*, we generated 8 different polynomials by repeatedly multiplying with x.

Taking the coefficients out, can be represented by using the fact that GF(3^2) is a vector space overZ_3. We consider a linear map as:
ϕ: GF(3^2)$\rightarrow Z_3$
We applied such a functional to obtain the desired sequence.
The interesting part is that if we chose another functional, as, the one that adds the coefficients of each polynomial, we would have arrived at

010122021

from which we can obtain a De Bruijn sequence by inserting 0 at the appropriate place

It seems, though, that we only generate all the elements of GF(p^l) in a particular order...but *how do we know they are de Bruijn sequences?*
Suppose that,

$$q(x) = x^l - a_{l-1}x^{l-1} - a_{l-2}x^{l-2} - ... - a_1 x - a_0$$

is irreducible over Z_p and that α=x is a primitive root of q(x) in GF(p^l)=Z_3/q(x)
Let ϕ: GF(p^l)$\rightarrow Z_p$ be any nonzero functional (linear when we view GF(p^l) as a vector space over Z_p)
Writing GF(p^l)*={α,α^2,...,α^{p^l-1}} and applying ϕ to its elements, we generate the sequence:

$$\phi(\alpha), \phi(\alpha^2), ..., \phi(\alpha^{p^l-1})$$

We assume that there is a repeat of length l after some point. Then, there exist i and j satisfying 1$\leq i < j \leq$p^l-1 such that:

$$\phi(\alpha^i) = \phi(\alpha^j)$$

$$\phi(\alpha^{i+1}) = \phi(\alpha^{j+1})$$

...

$$\phi(\alpha^{i+l-1}) = \phi(\alpha^{j+l-1})$$

This is like saying that the l number of vectors $\alpha^{j+k} - \alpha^{i+k}, k \in \{0, ...l-1\}$ are in the kernel of the linear functional ϕ
Since *dim(GF(p^l))*=l, then those vectors are linearly independent
So, there are constants $a_k \in$ Z_p, not all 0, for which

$$\sum_{k=0}^{l-1} a_k(\alpha^{j+k} - \alpha^{i+k}) = 0$$

or:

$$\alpha^i(1 - \alpha^{j-i}) \sum_{k=0}^{l-1} a_k \alpha^k = 0$$

We know $\alpha^i \neq 0$ and the only way to have 1=α^{j-i} is i=j, so it must be the case that the sum is equal to 0. However, this cannot happen, since then g(α)=0 for a polynomial g(x) of degree l-1, contradicting the assumption. So, no repeated windows are possible.

The reason we need to insert a zero into the sequence:

$$\phi(\alpha), \phi(\alpha^2), ..., \phi(\alpha^{p^l-1})$$

to produce a de Bruijn is that we will not generate 0 in GF(p^l) using this process

We used the multiplicative structure of GF(p^l) to generate different elements of GF(p^l) * in specific order and then use the linear map to get a Z$_p$ result.