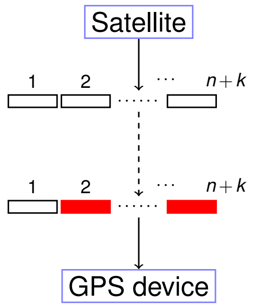
Erasure Codes.



n packet message. So send $n+k$!

Lose k packets.

Any n packets is enough!

n packet message.

Optimal.

Erasure Code: Example.

Send message of 1,4, and 4.

Make polynomial, $P(x) = a_2x^2 + a_1x + a_0$
with $P(1) = 1$, $P(2) = 4$, $P(3) = 4$.

How?
Lagrange Interpolation.
Linear System.

Work modulo 5.

$P(x) = x^2 \pmod 5$
$P(1) = 1, P(2) = 4, P(3) = 9 = 4 \pmod 5$

Send $(0, P(0)) \ldots (5, P(5))$.

6 points. Better work modulo 7 at least!

Why? $(0, P(0)) = (5, P(5)) \pmod 5$

Example

Make polynomial, $P(x) = a_2x^2 + a_1x + a_0$
with $P(1) = 1$, $P(2) = 4$, $P(3) = 4$.

Modulo 7 to accommodate at least 6 packets.

Linear equations:

$$\begin{aligned} P(1) = a_2 + a_1 + a_0 &\equiv 1 \pmod 7 \\ P(2) = 4a_2 + 2a_1 + a_0 &\equiv 4 \pmod 7 \\ P(3) = 2a_2 + 3a_1 + a_0 &\equiv 4 \pmod 7 \end{aligned}$$

$6a_1 + 3a_0 = 2 \pmod 7$, $5a_1 + 4a_0 = 0 \pmod 7$

$a_1 = 2a_0 \pmod 7$ $a_0 = 2 \pmod 7$ $a_1 = 4 \pmod 7$ $a_2 = 2 \pmod 7$

$P(x) = 2x^2 + 4x + 2$

$P(1) = 1$, $P(2) = 4$, and $P(3) = 4$

Send
Packets: $(1,1), (2,4), (3,4), (4,7), (5,2), (6,0)$

Notice that packets contain "x-values".

Bad reception!

Send: $(1,1), (2,4), (3,4), (4,7), (5,2), (6,0)$

Recieve: (1,1) (3,4), (6,0)
Reconstruct?

Format: $(i, R(i))$.

Lagrange or linear equations.

$$\begin{aligned} P(1) = a_2 + a_1 + a_0 &\equiv 1 \pmod 7 \\ P(2) = 4a_2 + 2a_1 + a_0 &\equiv 4 \pmod 7 \\ P(6) = 2a_2 + 3a_1 + a_0 &\equiv 0 \pmod 7 \end{aligned}$$

Channeling my inner linear algebra genius ...

$P(x) = 2x^2 + 4x + 2$

Message? $P(1) = 1, P(2) = 4, P(3) = 4$.

Questions for Review

You want to encode a secret consisting of 1,4,4.

How big should modulus be?
Larger than 144 and prime!

You want to send a message consisting of packets 1,4,2,3,0

through a noisy channel that loses 3 packets.

How big should modulus be?
Larger than 8 and prime!

Send n packets b-bit packets, with k errors.
Modulus should be larger than $n+k$ and also larger than 2^b.

Polynomials.

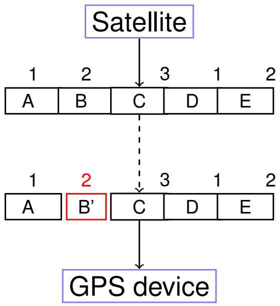
- ..give Secret Sharing.
- ..give Erasure Codes.

Error Correction:

Noisy Channel: corrupts k packets. (rather than loss.)

Additional Challenge: Finding which packets are corrupt.

Error Correction



3 packet message. Send 5.

Corrupts 1 packets.

Example.

Message: $3, 0, 6$.

Reed Solomon Code: $P(x) = x^2 + x + 1 \pmod 7$ has $P(1) = 3, P(2) = 0, P(3) = 6$ modulo 7.

Send: $P(1) = 3, P(2) = 0, P(3) = 6, P(4) = 0, P(5) = 3$.

(Aside: Message in plain text!)

Receive $R(1) = 3, R(2) = 1, R(3) = 6, R(4) = 0, R(5) = 3$.

$P(i) = R(i)$ for $n+k = 3+1 = 4$ points.

The Scheme.

Problem: Communicate n packets $m_1, \ldots, m_n$ on noisy channel that corrupts $\leq k$ packets.

Reed-Solomon Code:

1. Make a polynomial, $P(x)$ of degree $n-1$, that encodes message.
   - $P(1) = m_1, \ldots, P(n) = m_n$.
   - Comment: could encode with packets as coefficients.
2. Send $P(1), \ldots, P(n+2k)$.

After noisy channel: Recieve values $R(1), \ldots, R(n+2k)$.

Properties:
(1) $P(i) = R(i)$ for at least $n+k$ points i,
(2) $P(x)$ is unique degree $n-1$ polynomial that contains $\geq n+k$ received points.

Slow solution.

Brute Force:
For each subset of $n+k$ points
Fit degree $n-1$ polynomial, $Q(x)$, to n of them.
Check if consistent with $n+k$ of the total points.
If yes, output $Q(x)$.

- For subset of $n+k$ pts where $R(i) = P(i)$, method will reconstruct $P(x)$!
- For any subset of $n+k$ pts,
  1. there is unique degree $n-1$ polynomial $Q(x)$ that fits n of them
  2. and where $Q(x)$ is consistent with $n+k$ points
     $\implies P(x) = Q(x)$.

Reconstructs $P(x)$ and only $P(x)$!!

Properties: proof.

$P(x)$: degree $n-1$ polynomial.
Send $P(1), \ldots, P(n+2k)$
Receive $R(1), \ldots, R(n+2k)$
At most k i's where $P(i) \neq R(i)$.

Properties:
(1) $P(i) = R(i)$ for at least $n+k$ points i,
(2) $P(x)$ is unique degree $n-1$ polynomial that contains $\geq n+k$ received points.

Proof:
(1) Sure. Only k corruptions.
(2) Degree $n-1$ polynomial $Q(x)$ consistent with $n+k$ points.
$Q(x)$ agrees with $R(i)$, $n+k$ times.
$P(x)$ agrees with $R(i)$, $n+k$ times.

Total points contained by both:	$2n+2k$.	P	Pigeons.
Total points to choose from	: $n+2k$.	H	Holes.
Points contained by both	: $\geq n$.	$\geq P-H$	Collisions.

$\implies Q(i) = P(i)$ at n points.
$\implies Q(x) = P(x)$. □

Example.

Received $R(1) = 3, R(2) = 1, R(3) = 6, R(4) = 0, R(5) = 3$

Find $P(x) = p_2x^2 + p_1x + p_0$ that contains $n+k = 3+1$ points.

All equations..

$$\begin{aligned}
p_2 + p_1 + p_0 &\equiv 3 \pmod 7 \\
4p_2 + 2p_1 + p_0 &\equiv 1 \pmod 7 \\
2p_2 + 3p_1 + p_0 &\equiv 6 \pmod 7 \\
2p_2 + 4p_1 + p_0 &\equiv 0 \pmod 7 \\
1p_2 + 5p_1 + p_0 &\equiv 3 \pmod 7
\end{aligned}$$

Assume point 1 is wrong and solve..no consistent solution!
Assume point 2 is wrong and solve...consistent solution!

In general..

$P(x) = p_{n-1}x^{n-1} + \cdots p_0$ and receive $R(1), \ldots R(m = n+2k)$.

$$
\begin{aligned}
p_{n-1} + \cdots p_0 &\equiv R(1) \pmod p \\
p_{n-1}2^{n-1} + \cdots p_0 &\equiv R(2) \pmod p \\
&\;\cdot \\
p_{n-1}i^{n-1} + \cdots p_0 &\equiv R(i) \pmod p \\
&\;\cdot \\
p_{n-1}(m)^{n-1} + \cdots p_0 &\equiv R(m) \pmod p
\end{aligned}
$$

Error!! .... Where???
Could be anywhere!!! ...so try everywhere.
Runtime: $\binom{n+2k}{k}$ possibilitities.

Something like $(n/k)^k$ ...Exponential in k!.

How do we find where the bad packets are efficiently?!?!?!

Ditty...

Oh where, Oh where
has my little dog gone?
Oh where, oh where can he be

With his ears cut short
And his tail cut long
Oh where, oh where can he be?

Oh where, Oh where
have my packets gone.. wrong?
Oh where, oh where do they not fit.

With the polynomial well put
But the channel a bit wrong
Where, oh where do we look?

Where oh where can my bad packets be?

$$
\begin{aligned}
E(1)(p_{n-1} + \cdots p_0) &\equiv R(1)E(1) \pmod p \\
\mathbf{0} \times E(2)(p_{n-1}2^{n-1} + \cdots p_0) &\equiv R(2)E(2) \pmod p \\
&\;\vdots \\
E(m)(p_{n-1}(m)^{n-1} + \cdots p_0) &\equiv R(n+2k)E(m) \pmod p
\end{aligned}
$$

Idea: Multiply equation i by 0 if and only if $P(i) \neq R(i)$.
All equations satisfied!!!!!

But which equations should we multiply by 0? Where oh where...??

We will use a polynomial!!! That we don't know. But can find!

Errors at points $e_1, \ldots, e_k$. (In diagram above, $e_1 = 2$.)

Error locator polynomial: $E(x) = (x - e_1)(x - e_2) \ldots (x - e_k)$.

$E(i) = 0$ if and only if $e_j = i$ for some j

Multiply equations by $E(\cdot)$. (Above E(x) = (x-2).)

All equations satisfied!!

Example.

Received $R(1) = 3, R(2) = 1, R(3) = 6, R(4) = 0, R(5) = 3$

Find $P(x) = p_2x^2 + p_1x + p_0$ that contains $n + k = 3 + 1$ points.

Plugin points...

$$
\begin{aligned}
(1-2)(p_2 + p_1 + p_0) &\equiv (3)(1-2) \pmod 7 \\
(2-2)(4p_2 + 2p_1 + p_0) &\equiv (1)(2-2) \pmod 7 \\
(3-2)(2p_2 + 3p_1 + p_0) &\equiv (6)(3-2) \pmod 7 \\
(4-2)(2p_2 + 4p_1 + p_0) &\equiv (0)(4-2) \pmod 7 \\
(5-2)(4p_2 + 5p_1 + p_0) &\equiv (3)(5-2) \pmod 7
\end{aligned}
$$

Error locator polynomial: $(x-2)$.

Multiply equation i by $(i-2)$. All equations satisfied!

But don't know error locator polynomial! Do know form: $(x-e)$.

4 unknowns (p_0, p_1, p_2 and e), 5 nonlinear equations.

..turn their heads each day,

$$
\begin{aligned}
E(1)(p_{n-1} + \cdots p_0) &\equiv R(1)E(1) \pmod p \\
&\;\vdots \\
E(i)(p_{n-1}i^{n-1} + \cdots p_0) &\equiv R(i)E(i) \pmod p \\
&\;\vdots \\
E(m)(p_{n-1}(n+2k)^{n-1} + \cdots p_0) &\equiv R(m)E(m) \pmod p
\end{aligned}
$$

...so satisfied, I'm on my way.

$m = n+2k$ satisfied equations, $n+k$ unknowns. But nonlinear!

Let $Q(x) = E(x)P(x) = a_{n+k-1}x^{n+k-1} + \cdots a_0$.

Equations:

$$Q(i) = R(i)E(i).$$

and linear in a_i and coefficients of $E(x)$!

Finding $Q(x)$ and $E(x)$?

- $E(x)$ has degree k ...

$$E(x) = x^k + b_{k-1}x^{k-1} \cdots b_0.$$

$\implies$ k (unknown) coefficients. Leading coefficient is 1.

- $Q(x) = P(x)E(x)$ has degree $n+k-1$ ...

$$Q(x) = a_{n+k-1}x^{n+k-1} + a_{n+k-2}x^{n+k-2} + \cdots a_0$$

$\implies$ $n+k$ (unknown) coefficients.

Number of unknown coefficients: $n+2k$.

Solving for $Q(x)$ and $E(x)$...and $P(x)$

For all points $1,\ldots,i,n+2k=m$,

$$Q(i)=R(i)E(i) \pmod{p}$$

Gives $n+2k$ linear equations.

$$\begin{aligned}
a_{n+k-1}+\ldots a_0 &\equiv R(1)(1+b_{k-1}\cdots b_0) \pmod{p}\\
a_{n+k-1}(2)^{n+k-1}+\ldots a_0 &\equiv R(2)((2)^k+b_{k-1}(2)^{k-1}\cdots b_0) \pmod{p}\\
&\vdots\\
a_{n+k-1}(m)^{n+k-1}+\ldots a_0 &\equiv R(m)((m)^k+b_{k-1}(m)^{k-1}\cdots b_0) \pmod{p}
\end{aligned}$$

..and $n+2k$ unknown coefficients of $Q(x)$ and $E(x)$!

Solve for coefficients of $Q(x)$ and $E(x)$.

Find $P(x)=Q(x)/E(x)$.

Example.

Received $R(1)=3, R(2)=1, R(3)=6, R(4)=0, R(5)=3$

$Q(x)=E(x)P(x)=a_3x^3+a_2x^2+a_1x+a_0$

$E(x)=x-b_0$

$Q(i)=R(i)E(i)$.

$$\begin{aligned}
a_3+a_2+a_1+a_0 &\equiv 3(1-b_0) \pmod{7}\\
a_3+4a_2+2a_1+a_0 &\equiv 1(2-b_0) \pmod{7}\\
6a_3+2a_2+3a_1+a_0 &\equiv 6(3-b_0) \pmod{7}\\
a_3+2a_2+4a_1+a_0 &\equiv 0(4-b_0) \pmod{7}\\
6a_3+4a_2+5a_1+a_0 &\equiv 3(5-b_0) \pmod{7}
\end{aligned}$$

$a_3=1$, $a_2=6$, $a_1=6$, $a_0=5$ and $b_0=2$.

$Q(x)=x^3+6x^2+6x+5$.

$E(x)=x-2$.

Example: finishing up.

$Q(x)=x^3+6x^2+6x+5$.
$E(x)=x-2$.

```
                1 x^2 + 1 x + 1
        ----------------
x - 2 ) x^3  + 6 x^2 + 6 x + 5
        x^3  - 2 x^2
        ----------
               1 x^2 + 6 x + 5
               1 x^2 - 2 x
               --------------
                       x + 5
                       x - 2
                       -----
                           0
```

$P(x)=x^2+x+1$

Message is $P(1)=3, P(2)=0, P(3)=6$.

What is $\frac{x-2}{x-2}$? 1

Except at $x=2$? Hole there?

Error Correction: Berlekamp-Welsh

Message: $m_1,\ldots,m_n$.

Sender:

1. Form degree $n-1$ polynomial $P(x)$ where $P(i)=m_i$.
2. Send $P(1),\ldots,P(n+2k)$.

Receiver:

1. Receive $R(1),\ldots,R(n+2k)$.
2. Solve $n+2k$ equations, $Q(i)=E(i)R(i)$ to find $Q(x)=E(x)P(x)$ and $E(x)$.
3. Compute $P(x)=Q(x)/E(x)$.
4. Compute $P(1),\ldots,P(n)$.

Check your undersanding.

You have error locator polynomial!

Where oh where have my packets gone wrong?

Factor? Sure.
Check all values? Sure.

Efficiency? Sure. Only $n+k$ values.
See where it is 0.

Hmmm...

Is there one and only one $P(x)$ from Berlekamp-Welsh procedure?

Existence: there is a $P(x)$ and $E(x)$ that satisfy equations.

Unique solution for $P(x)$

Uniqueness: any solution $Q'(x)$ and $E'(x)$ have

$$\frac{Q'(x)}{E'(x)} = \frac{Q(x)}{E(x)} = P(x). \quad (1)$$

Proof:
We claim

$$Q'(x)E(x) = Q(x)E'(x) \text{ on } n+2k \text{ values of } x. \quad (2)$$

Equation **??** implies **??**:

$Q'(x)E(x)$ and $Q(x)E'(x)$ are degree $n+2k-1$
and agree on $n+2k$ points
$E(x)$ and $E'(x)$ have at most k zeros each.
Can cross divide at n points.
$\implies \frac{Q'(x)}{E'(x)} = \frac{Q(x)}{E(x)}$ equal on n points.
Both degree $\leq n \implies$ Same polynomial! □

Last bit.

Fact: $Q'(x)E(x) = Q(x)E'(x)$ on $n+2k$ values of x.

Proof: Construction implies that

$$Q(i) = R(i)E(i)$$
$$Q'(i) = R(i)E'(i)$$

for $i \in \{1, \ldots n+2k\}$.

If $E(i) = 0$, then $Q(i) = 0$. If $E'(i) = 0$, then $Q'(i) = 0$.
$\implies Q(i)E'(i) = Q'(i)E(i)$ holds when $E(i)$ or $E'(i)$ are zero.

When $E'(i)$ and $E(i)$ are not zero

$$\frac{Q'(i)}{E'(i)} = \frac{Q(i)}{E(i)} = R(i).$$

Cross multiplying gives equality in fact for these points. □

Points to polynomials, have to deal with zeros!

Example: dealing with $\frac{x-2}{x-2}$ at $x = 2$.

Berlekamp-Welsh algorithm decodes correctly when k errors!

Summary. Error Correction.

Communicate n packets, with k erasures.

How many packets? $n+k$
How to encode? With polynomial, $P(x)$.
Of degree? $n-1$
Recover? Reconstruct $P(x)$ with any n points!

Communicate n packets, with k errors.

How many packets? $n+2k$
Why?
k changes to make diff. messages overlap
How to encode? With polynomial, $P(x)$. Of degree? $n-1$.
Recover?
Reconstruct error polynomial, $E(X)$, and $P(x)$!
Nonlinear equations.
Reconstruct $E(x)$ and $Q(x) = E(x)P(x)$. Linear Equations.
Polynomial division! $P(x) = Q(x)/E(x)$!

Reed-Solomon codes. Welsh-Berlekamp Decoding. Perfection!

Wow.

Lots of material today...