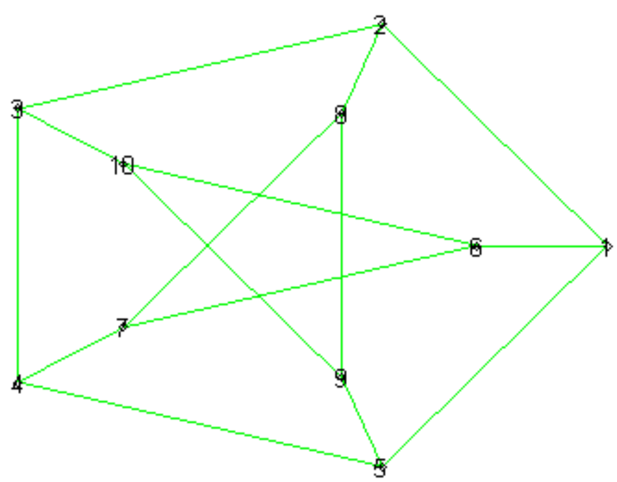
Commutative Algebra Meets Network Reliability

Will Traves, U.S. Naval Academy

April 05, 2001

Math Colloquium, Towson University



Commutative Algebra $\longleftrightarrow$ Algebraic Geometry

$\downarrow$

Combinatorics of Networks

Networks arise in daily life, in scientific models and even in our recreational pursuits.

• August 10, 1996: A fault in two power lines in Oregon led to blackouts in 11 US states

• May 04, 2000: Love Bug Virus spreads over the internet, causing billions of dollars in damage

• Recreational Examples: Six Degrees of Separation
    • Kevin Bacon and Marlon Brando (Die Zeit)
    • Paul Erdős
    • Monika Lewinsky (NY Times, 21 Feb. 1998)

• Mathematical Genealogy Project:
http://hcoonce.math.mankato.msus.edu/

• Scientific applications: topology of food webs, electrical power grids, metabolic networks, the WWW, the internet backbone, telephone call networks, etc. (S. Strogatz, Nature, March 2001)

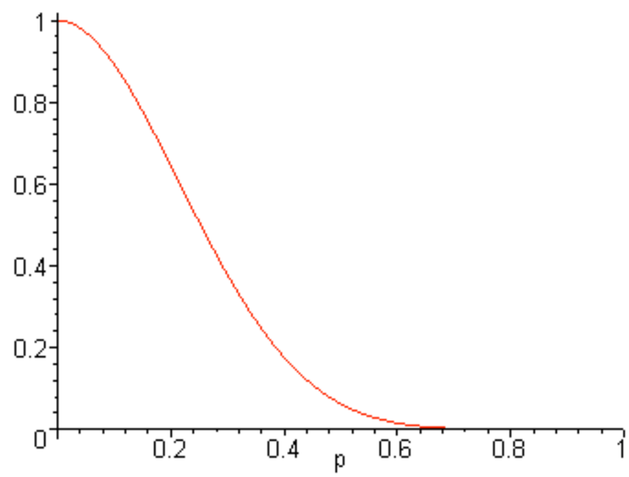
Tools from Commutative Algebra allow us to estimate the reliability of networks like the internet backbone or a small intranet.

Graph: G
Number of vertices of G**:** n
Number of edges of G**:** m
We allow multiple edges between vertices and loops.

We say that G **is connected when any two vertices are connected by a path.**

Let p **represent the probability that any given edge fails (is removed). We are interested in the reliability function,**

$$\mathrm{Rel}_G(p) = \Pr(G \textit{ remains connected}).$$



In general, it is too difficult to compute the reliability function exactly. This leads to our central problem:

Bound the reliability function in terms of information that can be easily computed from the graph (say, in polynomial time).

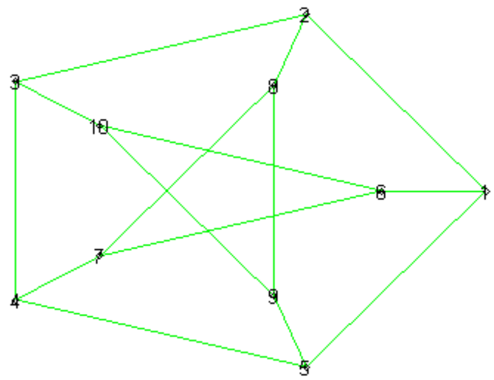
Plan of Attack:

- $\mathrm{Rel}_G(p)$ **is a polynomial in** p
- **Using connections between simplicial complexes and commutative algebra, associate a ring** R_G **to the graph** G
- **The Hilbert polynomial of** R_G **measures the structure of** R_G **and is related to the reliability polynomial,** $\mathrm{Rel}_G(p)$**.**
- **The ring** R_G **is Cohen-Macaulay and this puts constraints on its Hilbert polynomial**
- **These constraints lead to tight bounds on the reliability function.**

Provan and Ball. ***Bounds on the reliability polynomial for shellable independence systems.*** **SIAM J. Alg. Disc. Meth. Volume 3, No. 2, 1982.**

$\text{Rel}_G(p)$ is a polynomial in p

To illustrate, consider Petersen's Graph:



$n = 10, m = 15$

We need to remove at least 3 edges to disconnect G (cut number k=3)

Removing <u>any</u> 7 edges disconnects G (dimension d=6)

Removing as many edges as possible while preserving connectivity leads to a spanning tree.

spanning: connected subgraph that involves all vertices of G

tree: no cycles (paths that are loops)

All spanning trees have $n-1\,(=9)$ edges.

The probability of any given connected spanning graph occurring in which *i* edges have failed is

$$p^i(1-p)^{n-i}.$$

Let f_i be the number of ways to remove *i* edges from G while preserving connectivity. Then f_i is the number of spanning graphs with $n-i$ edges.

$$\mathrm{Rel}_G(p) = \sum_{i=0}^{d} f_i p^i (1-p)^{n-i}.$$

Easy to compute $f_0, f_1, \ldots, f_{k-1}$ (binomial coefficients)

To compute f_d, use the Matrix-Tree Theorem (Kirchoff): **A: adjacency matrix of G**
D: degree matrix (diagonal)
|det(cofactor(A-D))| = number of spanning trees

$$\det\left(\begin{bmatrix} -3 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -3 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & -3 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -3 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -3 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & -3 \end{bmatrix}\right) = 2000$$

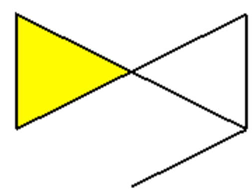
Not too difficult to compute f_k
(Here: $f_0 = 1, f_1 = 15, f_2 = 105, f_3 = 445, f_d = 2000.$)

A Simplicial Complex

$$\Delta_G = \{ sets\ S\ of\ edges : G - S\ remains\ connected \}$$

This is a simplicial complex because:
(1) $\phi \in \Delta_G$ **(2)** Δ_G **is closed under inclusion**

Note that f_i **is the number of sets of size *i* in** Δ_G



The Kruskal-Katona Theorem bounds the face numbers f_i **of our simplicial complex.**

Our complex has a special structure that admits an algebraic description.

Let I_Δ **be the ideal in** $k[x_1, \ldots, x_m]$ **generated by monomials whose support does not lie in** Δ_G

<u>Ex</u>: $x_1 x_2 x_4^2 \in I_\Delta$ **: its support** $\{1,2,4\} \notin \Delta_G$

Stanley-Reisner ring: $R_G = k[x_1, \ldots, x_m] / I_\Delta$

After "killing" I_Δ **the only monomials that remain have support in the simplicial complex** Δ_G**.**

The ring R_G is a graded ring. That is, it decomposes into pieces, each of which corresponds to homogeneous polynomials of a certain degree:

$$R_G = k \oplus R_1 \oplus R_2 \oplus \cdots$$

(This grading exists because the ideal we are killing is generated by homogeneous polynomials – in fact, by monomials)

The Hilbert Series $Hilb(R_G, t)$ is the generating function of the dimensions of the graded pieces:

$$Hilb(R_G, t) = \sum_{n=0}^{\infty} (\dim_k R_n) t^n$$

<u>Ex</u>: Let $R = k[x_1, x_2, \ldots, x_d]$. Then $\dim_k(R_n) = \binom{d+n-1}{d-1}$

so $$Hilb(k[x_1, x_2, \ldots, x_d], t) = \sum_{n=0}^{\infty} \binom{d+n-1}{d-1} t^n = \frac{1}{(1-t)^d}.$$

Here $I_\Delta = (0)$ and $G = 2^d$, the complete complex on d vertices. The dimension of the complex is d and the dimension of the ring is the order of the pole at t=1 in the Hilbert Series (also d). In general,

$$\dim R_G = \dim \Delta_G.$$

Theorem (Hilbert-Serre): $Hilb(R_G, t) = h_G(t) \big/ (1-t)^d$

What does the Hilbert Series look like for R_G?

Theorem: $$Hilb(R_G, t) = \sum_{i=0}^{d} f_i \left(\frac{t}{1-t}\right)^i$$

Proof: We induce a "fine grading" on R_G **by assigning degree** $(0,\ldots,0,1^i,0,\ldots,0)$ **to** x_i**. The monomials clearly generate** R_G **as a k-vector space, but only those with support in** Δ_G **are nonzero.**

$$Hilb(R_G, t_1, \ldots, t_m) = \sum_{sppt(a)\in\Delta} t^a$$

$$= \sum_{F\in\Delta} \sum_{sppt(a)=F} t^a$$

$$= \sum_{F\in\Delta} \prod_{v\in F} \left(\frac{t_v}{1-t_v}\right) \quad (*)$$

This last equation needs some justification:

First note that $\frac{t_v}{1-t_v} = t_v(1 + t_v + t_v^2 + \cdots) = t_v + t_v^2 + t_v^3 + \cdots$

so that

$$\prod_v \frac{t_v}{1-t_v} = \prod_v (t_v + t_v^2 + t_v^3 + \cdots) = \sum \mathit{monomials\ involving\ all}\ t_v\text{'s}.$$

Now set all $t_v = t$ **in (*) to get**

$$Hilb(R_G, t) = \sum_{i=0}^{d} f_i \left(\frac{t}{1-t}\right)^i.$$

Theorem: $$Hilb(R_G,t) = \sum_{i=0}^{d} f_i \left(\frac{t}{1-t}\right)^i$$

Theorem: $$Hilb(R_G,t) = h_G(t) \big/ (1-t)^d$$

$$\boxed{\begin{aligned}(1-t)^{m-d} h_G(t) &= (1-t)^m Hilb(R_G,t)\\ &= \sum_{i=0}^{d} f_i t^i (1-t)^{m-i}\\ &= \mathrm{Rel}_G(t).\end{aligned}}$$

So the reliability polynomial of G is determined by the Hilbert polynomial of R_G.

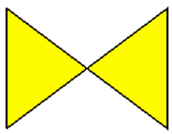
The coefficients of the Hilbert polynomial $h_G(p) = h_0 + h_1 p + \cdots + h_d p^d$ are related to the face numbers:

$$h_k = \sum_{i=0}^{k} (-1)^{k-i} \binom{d-i}{k-i} f_i \qquad \sum_{i=0}^{d} h_i = f_d$$

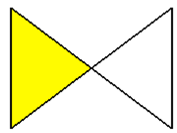
For instance, in our running example,

$$h_0 = 1, \quad h_1 = 9, \quad h_2 = 45, \quad h_3 = 155.$$

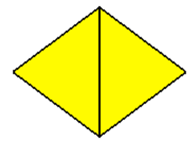
The simplicial complex Δ_G is pure: every maximal set in Δ_G has the same size, d



pure



not pure



strongly pure

In fact, much more is true: every induced subsimplicial complex of Δ_G is pure. (In combinatorics we would say that Δ_G is a matroid). This corresponds to a strong equidimensionality result on R_G, sufficient to ensure that R_G is Cohen-Macaulay.

Now Macaulay's Theorem puts severe constraints on the coefficients of the Hilbert polynomial of R_G

Macaulay's Theorem: $0 \le h_j \le h_i^{<j/i>}$ **for all** $0 \le i < j \le d$

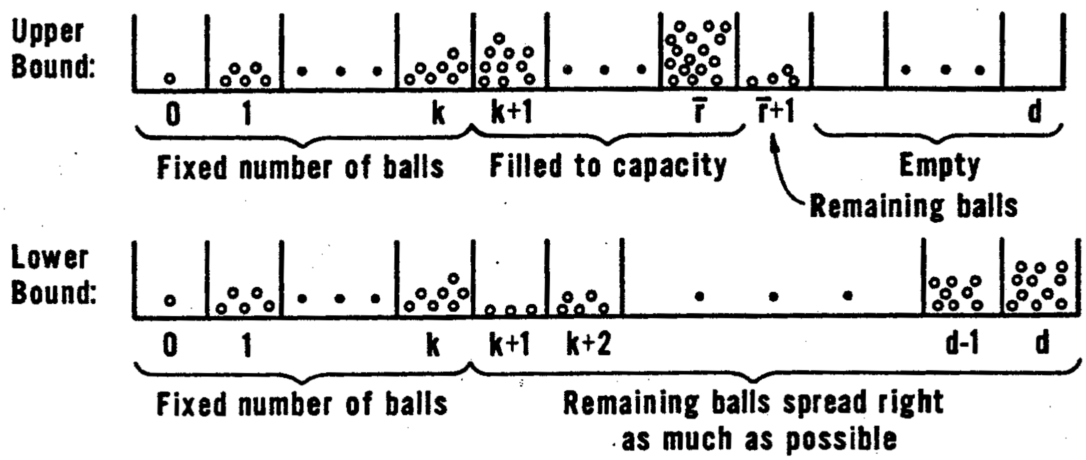
<u>Ex</u>: (Petersen's Graph)

$$h_3 = 155 = \binom{10}{3} + \binom{8}{2} + \binom{7}{1} \qquad h_4 \le h_3^{<4/3>} = \binom{11}{4} + \binom{9}{3} + \binom{8}{2} = 442$$

Similarly: $$h_5 \le 1086, \quad h_6 \le 2388.$$

Constructing Bounds on the Hilbert Polynomial

Since $h_j \geq 0$ and $\sum_{i=0}^{d} h_i = f_d$ we can interpret assigning values to the coefficients h_j as placing f_d balls into $d+1$ boxes.



Since $0 \leq p \leq 1$ the polynomial $h_G(p) = h_0 + h_1 p + \cdots + h_d p^d$ is largest when the first free coefficients are as big as possible and is smallest when the last free coefficients are as big as possible.

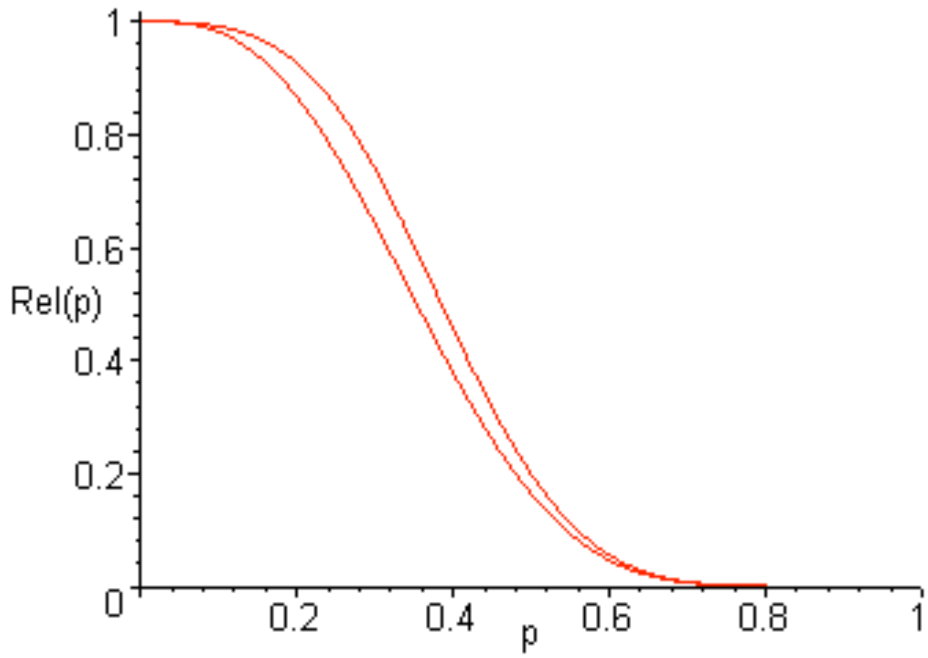
In our example, an upper bound is obtained by putting as many balls as possible (subject to Macaulay's Theorem) into the lowest numbered boxes, until we run out of balls. This leads to the upper bound:

$$h_G(p) = 1 + 9p + 45p^2 + 155p^3 + 442p^4 + 1086p^5 + 262p^6$$

Applying Macaulay's Theorem to the lower bound is more difficult: in order to put any ball in the box *d*, we must put some in box *d-1*, etc. There is a formula for these numbers, leading to the lower bound:

$$h_G(p) = 1 + 9p + 45p^2 + 155p^3 + 243p^4 + 523p^5 + 1024p^6$$

Results for Petersen's Graph



Upper bound on the reliability polynomial:

$$(1-p)^9(1+9p+45p^2+155p^3+442p^4+1086p^5+262p^6)$$

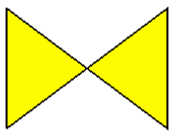
Lower bound on the reliability polynomial:

$$(1-p)^9(1+9p+45p^2+155p^3+243p^4+523p^5+1024p^6)$$

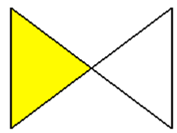
When the edge failure probability is 0.1, there is about a 98% probability that the graph remains connected.

When the edge failure probability is 0.05, there is about a 99.8% probability that the graph remains connected.

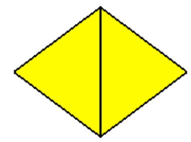
The simplicial complex Δ_G is pure: every maximal set in Δ_G has the same size, d



pure



not pure



strongly pure

In fact, much more is true: every induced subsimplicial complex of Δ_G is pure. (In combinatorics we would say that Δ_G is a matroid).

<u>Proof</u>: Suppose that we build a subsimplicial complex of Δ_G on all vertices save those in the set S. The faces in the subcomplex are collections of edges disjoint from S whose removal preserves the connectivity of G. Since we are only interested in connectivity, we can identify those vertices in G connected by edges in S (and remove all edges in S from the resulting graph H). Then the subcomplex is just Δ_H, and the result follows from the fact that that maximal faces in the subcomplex correspond to spanning trees in H all of which have the same size (number of vertices – 1).