# Cybernetics in Analysis and Synthesis of Networked Systems

Heikki Hyötyniemi

Cybernetics Group
Helsinki University of Technology, Control Engineering Laboratory
P.O. Box 5500, FIN-02015 HUT, Finland
heikki.hyotyniemi@hut.fi
http://www.control.hut.fi/hyotyniemi

**Abstract.** It seems that the cybernetic system framework is well suited for analysis and design of technical systems. This approach offers useful models for qualitative characterization of fractal systems, and for design of such networks. As an example, a distributed energy production network is studied. There are interesting intuitions about what might be the nature of the robustness in the life-like networks.

## 1 Introduction

Today, there is great need for analysis and design methods for different kinds of networks and distributed agent systems. This is clear in environments like Internet; but also in social systems, for example, one would need tools for analysing and constructing networks where individual actors have differing ability profiles. Today's agent systems consist of software architectures with no underlying systemic theories. The resulting control schemes are, after all, centrally controlled rather than truly distributed; all actors contribute dorectly to the same goal. Indeed, the only added value in such systems is distribution of the work load.

*Complexity theory* is one of those frameworks where complex networks are being studied, and *emergence* is the keyword for reaching the levels of higher understanding [1]. The properties of complex systems have been studied within this framework using names like *self-organised criticality, phase transitions, edge of chaos,* and *highly optimised tolerance* [4]. The promises are huge, but there still exist very few practical engineering tools for analysis or synthesis based on these ideas.

It may still be that the route to attacking holistic phenomena is through reductionistic approaches. New mathematical results in the theory of *cybernetic systems* promise that distribution of control can result in emergent, unanticipated, theoretically interesting behaviors, and quantitative analysis and synthesis tools may now be available. To illustrate this, a concrete example from the field of *distributed energy production* is presented here from different points of view: First, the statistical properties of such systems are studied in general terms — indeed, an extension of the cybernetic models towards *multiplicative*

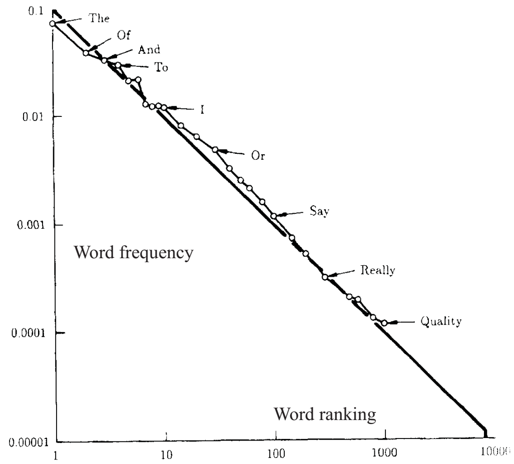


**Fig. 1.** Word frequencies are distributed according to the Zipf Law

*systems* is presented. After that, a system where the distribution is carried out in a cybernetically motivated manner is presented and its properties are studied.

It has been claimed that natural systems are more robust than the man-made ones. For example, a single fault can result in a domino effect in an energy supply system, but single collapses of some individual species do not escalate into ecocatastrophes. But what is this "natural robustness" in the first place? The power plant case simulations offer interesting intuitions.

For more information on the theoretical properties, and to see other kinds of applications of *neocybernetics,* see [12].

## 2 Analysis of existing systems

There is a huge need for tools for analysis of complex networks. The problem is is that reductionistic approaches cannot explain the system-level phenomena — but are there any holistic approaches available?

### 2.1 Characterizing networks

When trying to attack a complex system starting from top, abstracting out all details and ingnoring the underlying realm, forgetting about the properties of specific domains, one has to employ some general principles that hopefully hold. There are not so many principles that could be proposed. Such intuitively appealing ideas are those of *self-similarity, fractality* and *scale indepedence.* It has been claimed that fractality pops up in all *scale-free* structures and networks (see [2], [3]).

Fractal dimension $D$ is an extension of the dimensionality concept towards fractional values, and it is defined as:

$$D = \frac{\log N}{\log S}. \qquad (1)$$

Here, $N$ stands for the number of self-similar substructures, and $S$ represents the *scaling factor.* That is, assuming that a structure can be reconstructed using various smaller but equally shaped substructures, and assuming that this self-similarity continues at different scales, the above definition can be applied for determinig the dimension. Self-similarity and fractality has been observed in many natural forms – indeed, it has been said that fractality is the Geometry of Nature.

What are the consequences of this self-similarity? When the expression (1) is written in another form

$$\log N = D \log S, \tag{2}$$

it turns out that for a self-similar structure, the relationship between the quantities $S$ and $N$, when plotted in a log-log scale, is linear, the slope being determined by the fractal dimension. This *power law* or *Zipf law* pops up in seemingly very odd places: It had been noticed that city sizes within a country, word frequencies in a language, etc., seem to follow this law (see Figs. 1 and 2). The power law behaviour makes it easy to estimate the fractality beyond complex systems, and it makes it possible to estimate some of their properties on a higher, more abstract level.

What will be specially concentrated on in what follows, is technical systems, and specially *energy production networks.* Also in such systems, the power law seems to work fine — in Fig. 3, it is shown how frequently major power outages have been taking place.

The above figures are convincing, are they not? However, here it applies that "You See What You Expect to See": Are the dependencies really linear?!

### 2.2 Cybernetic networks

Why should a distributed energy network, for example, be strictly scale independent? Indeed, it is difficult to find natural systems that would remain completely invariant when the focus area is zoomed: The underlying principles, actors, and processes are so different in different levels that — even though the network structure remains — the connectedness of the nodes changes, thus ruining the scale-freeness. And if the fractal dimension changes as the scale changes, the linearity on the log-log scale does not hold. Indeed, it can be claimed that this scale independence is just an assumption that has been made to reach *something* concrete.

Are there any other general approaches to reaching something quantitative out from the very complex domain areas? Here it is claimed that a better applicable modeling principle is offered by *cybernetic* considerations.

A cybernetic system is a distributed system where there is no centralized control: The system level behavior is emergent, being result of interactions and feedbacks among the actors, or agents, in the network. As explained in [6], the typical property in a cybernetic system is its *strive towards balance.* All organizations reflect some balance, and all actions reflect strive towards better balance.

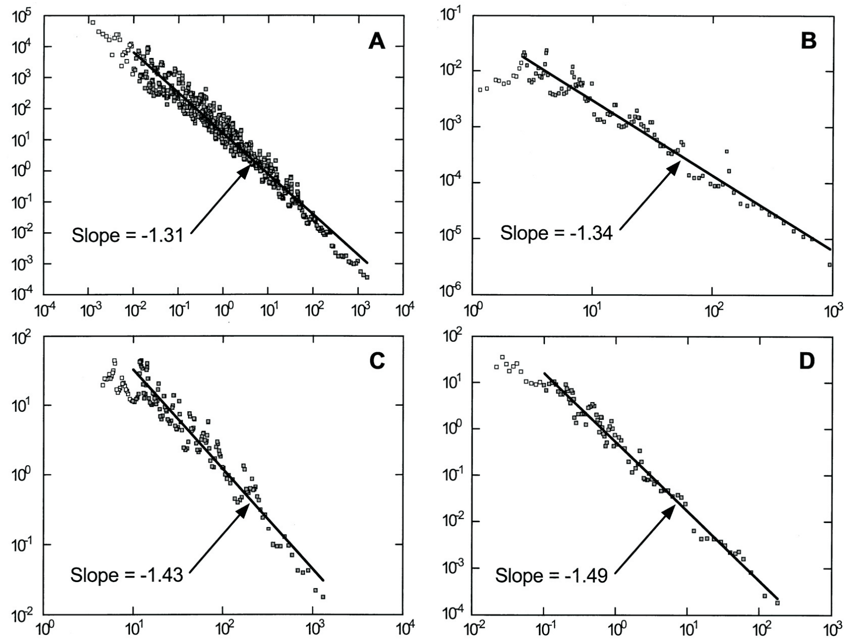


**Fig. 2.** Properties of forest fires [7]: (A) 4284 fires on U.S. Fish and Wildlife Service lands (1986-1995), (B) 120 fires in the Western United States (1950-1960), (C) 164 fires in Alaskan boreal forests (1990-1991), and (D) 298 fires in Australia (1926-1991). The number of fires is given as a function of the burnt area

Where do the balances come from, and why are they so characteristic to cybernetic systems?

In principle, a complete technical system with no possibility of failures could be constructed — but its price would be infinite. One always has to compromise: In larger technical systems there is always feedback from the economic reality. The system has to be implemented savingly, and when this saving-oriented thinking is extended to each level of the design, one already has a balance system. And cybernetics studies such balances among opposite needs where harmony is reached through opposing tensions (as Heraclitus would put it). In Fig. 4, a simple network of dependencies is shown. It is shown how the extent of power outages can also be seen as a variable in a cybernetic balance system: The company tries to keep things running, balancing between external and internal pressures. The customers should be kept satisfied, but as economically as possible; the difficulties are caused by the environmental conditions, thunder storms, etc., that cannot be controlled, and they must be taken into account as independent input variables. The graph in Fig. 4 can thus be seen as a realization of a cybernetic system behavior, striving towards dynamic balance (dotted lines denoting negative "forces", solid lines positive ones).

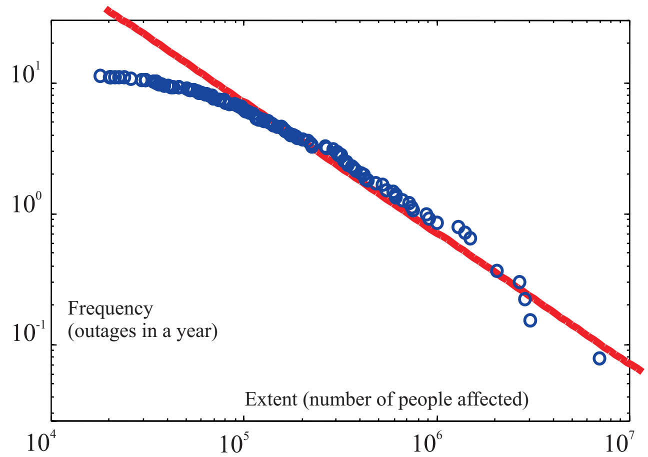


**Fig. 3.** Power outages in the US 1984–1997 (by John Doyle)

Existing complex systems, like distributed energy networks, have not knowingly been designed to be cybernetic. But because of the inevitable compromises, interactive feedback structures have been anyway integrated there — it is not a once-for-all design, but one typically has to reconstruct and fix the system as it is running, not to mention the continuous tuning of the production parameters. Learning organizations (also systems consisting of humans) react to their external and internal environments (see [10]). One could say that typically a "self-organized" system (even though humans may be needed to implement the self-organization) is cybernetic.

There is some evidence that seems to support the above analyses. In [5], similar considerations were studied from the practical point of view. A very complicated dynamic model was constructed there; disturbances were simulated, and outages resulted in correcting and preventive actions. Again, there was balancing between keeping the management satisfied and keeping the customers satisfied. The simulation results were qualitatively similar to the long-term outage statistics. In the paper, the "sandpile model" of cascading failures is used as a model, and it is claimed that the network system finds its balance near the Edge of Chaos; this is called Self-Organized Criticality (SOC).

Following the discussions in [6], a mathematically more compact representation of a cybernetic system can be given. It is assumed that the cybernetic essence can be captured in a *linear, dynamic, state-space model structure.* Even though the overall behavior can be very complex, locally, within a narrow regime around the nominal operating point, the internal system dynamics can be assumed to be locally linearizable:

$$\frac{d\,x}{d\,t}(t) = -\Lambda(x) A\, x(t) + \Lambda(x) B\, u, \tag{3}$$

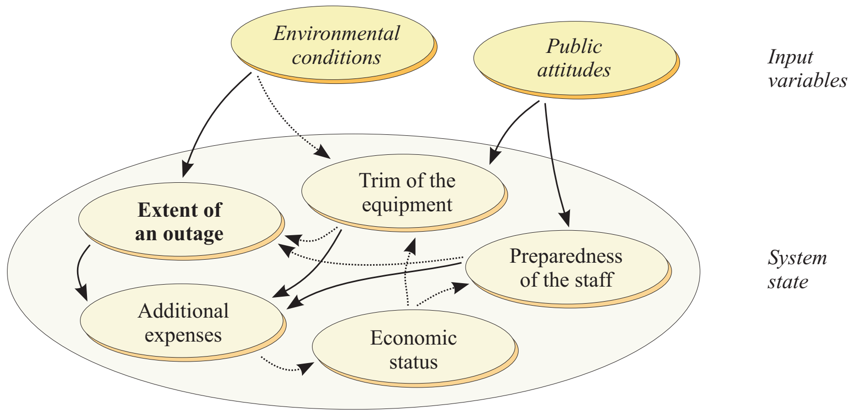


**Fig. 4.** How a system searches its balance

where the vector $x$ contains the effective cybernetic variables, and $u$ is the vector of environmental conditions, or inputs. The cybernetic variables are the system outputs, simultaneously used for feedback purposes. Matrices $A$ and $B$ determine the interactions among variables, dictating the main behavioral patterns of the system, and matrix $\Lambda(x)$ is an invertible matrix affecting the adaptation process. The main thing in the cybernetic system is its balance; assuming invertibility of $A$, this equilibrium can be solved as

$$\bar{x} = A^{-1} B\, u = \phi^T u. \tag{4}$$

It needs to be recognized that the above model is more or less hypothetical, because the elements in the matrices $A$ and $B$ are typically never explicitly known; the key point is that *it can be assumed that such an underlying model exists.* It does not matter if it is a mental model only, reflecting the more or less explicit company strategies: Anyway, reactions to environmental disturbances are compensated according to such a model by the operating company management. Because balances are only of interest, all dynamic parameters are not needed (compare to [5]).

Dynamics of $x$ is fast, meaning that the system immediately responds to current challenges, trying to survive in everyday life as there are changes in the environment $u$. The system matrices $A$ and $B$ are also adapted (see later), but slower, meaning that strategic planning takes place in the company, and the system is made more streamlined or otherwise better suited to its environment through structure adaptation.

When all subsystems are cybernetically balanced, can it be claimed that the system as a whole operates optimally? No, in the strictly cybernetic sense, or seen from the point of view of "higher-order cybernetics" (see [6]), it can be assumed that this kind of absolute optimum is never reached — there simply are so many alternative parameters that in practice cannot be exactly tuned. But even though the system may not be cybernetic in the higher-order sense, it is still “first-order cybernetic”: It is enough it works *somehow,* being capable of reacting to disturbances, and going towards (new) balances after them. A cybernetically suboptimal system still can be modeled applying the same basic structure (3).

It can be questioned whether such a simple model as that in (3) is really justifiable when describing complex economical decisions. There is a dilemma here: However uncertain and unquantifiable the phenomena and their dependencies are in the system, they are being routinely estimated and exploited in economical decisions today. And, truly, the interaction models — if they are employed in the first place — typically contain no fancy functions. There are formalized methodologies for assessing and comparing incompatible qualities (for example, see [9]), and there are risk theories, but they are typically based on simple one-issue-at-a-time comparisons and assessments, and no tools exist for systemic considerations. Indeed, looking at the modern decision support aids, it can be claimed that (3) is the most general model there exists.

In a complex world, there are no uniquely correct decisions. It is like with “parallel universes”: Different decisions result in different behaviors, and being an integral part of the environment, the world also changes accordingly. The resulting system with the modified structure is a balance system again (assuming that no pathological decisions are made), but the balance is different. As the earlier state of the world never comes back, nobody can say what would have happened if other actions would have been taken. Being a strategist being responsible for the decisions, you just need to “keep the tie tight” and convince the financers: It does not matter if you do not know the exact truths — nobody will ever know!

Even though the system structure were never explicitly formulated, the matrices can be re-engineered in the inverse way: If there is some observed behavior, the matrices can be reconstructed afterwards — in principle. In practice, there is too little data for concrete parameter identification. This applies specially to the dynamic model, because in a model structure (3) there are much more free parameters than there are in the static model matrix $\varphi$ in (4). The dynamic model is needed as a starting point to remind of the underlying realm of dynamic equilibria, but from the pragmatic point of view, the static balance model (4) is only studied. This can be expressed also in another way: We are concentrating on the *pattern view* rather than the *process view* when studying the complex system [11]. When characterizing a system, the relevant thing is where the system is eventually trying to get, not the actual route. The complex details of the nonlinear adaptation processes can be ignored if the final outcome is studied directly.

Summarizing the above discussion — it can be claimed that schematic illustrations as the one shown in Fig. 4, as figured out by the domain area expert, truly reflect the underlying system structure and the relevant dependencies. Even though the system matrices in (3) may remain unknown, the model structure helps to see some qualitative consequences. If the system is cybernetic, its “nature” is reflected in the balances.

To reach some more concrete results in addition to mere intuitions, it turns out that without numerical values not very much can be done. As compared to other applications in [12], where one is studying concrete populations, etc., the domain field is now much more abstract. How to quantify penomena, how to make variables compatible, and how to make the models truly functional?

### 2.3 "Bayesian cybernetics"

It is often difficult to quantify variables in a complex system; for example, how to numerically represent "extent of outage", "economic status", etc., in Fig. 4?

First, it can be noted that such variables are not absolute numbers; typically they are relative figures, proportional quantities, being contrasted and measured in terms of some optimal or maximal case. A handy framework involving such scaled variables is using probabilities (or possibilities): How probably an event will happen? If all variables are interpreted as probability values, one has a network of probabilities. Such networks have been studied a lot lately, mainly in the framework of *Bayesian networks* (for the original contribution, see [8]).

Bayesian networks are theoretically well-founded, based on probability theory, and reasoning applications can readily be implemented on such platforms. However, this well-foundedness only applies if the underlying assumptions are met: Nodes in the network that are not connected have to be independent of each other. Explicit dependency structures are denoted by arrows; problems emerge immediately if there are loops in the network topology. It is evident that in cybernetic systems where all variables are assumed to be in interaction one will have practical problems if trying to implement a Bayesian network model. If all variables are connected to each other, the probabilistic network can be maximally dense.

Indeed, it can be assumed that graphs like that in Fig. 4 characterize fully connected probabilistic networks, describing coupled phenomena. Cybernetic systems can be seen as "probability balance" systems. As compared to Bayesian networks, it can be seen that this approach nicely compensates the deficiencies of them. However, it needs to be noted that the resulting system is (as will be seen later) not Bayesian, as different kinds of formulas are applied; the framework is not even strictly probabilistic. This issue needs to be elaborated on.

To proceed, one needs to define variables that are more flexible than strictly probabilistic variables are. It seems that some new interpretations are motivated here. Let *relevance* $r_i > 0$ denote how relevant an event $i$ is; to some extent, relevances are assumed to have the same intuitive interpretation and manipulation rules as probabilities have. The *nominal* value of a relevance variable is 1, but the relevance values can exceed this, if the corresponding event is somehow more acute than in normal cases. Two relevance variables can be related to each other as

$$r_i \sim r_j^{a_{ij}}, \qquad (5)$$

where the parameter $a_{ij}$ represents the mutual contribution: To which extent a variable "belongs" to the other, or how well it explains the other. Essentially, the variables represent *fuzzy subsets,* and the contribution parameter reveals how near each other these subsets are. Further, there can be various variables — in such a case, their contributions are directly multiplied:

$$r_i \sim r_j^{a_{ij}} r_k^{a_{ik}}. \tag{6}$$

What if the relevances $r_j$ and $r_k$ are not independent? Can the mappings be inverted? — There are many questions that cannot be answered now. One should not look at these relevance issues as presented here from the narrow viewpoint of traditional probability theory: Indeed, now the whole thinking needs to be turned upside down. One should not be studying individual variables, but the whole cybernetic system; the goal is determine the relationships reflected in the balanced relevance values, and the looking at the results one can determine whether the new concepts are relevant in the first place. This will be elaborated on in what follows.

### 2.4 Multiplicative systems

If the variables are based on probabilities, or relevances, the linear functions are not applicable any more — the variables cannot be directly added together. Probabilities of independent variables typically have to be *multiplied* to reach meaningful results. For example, the overall risk for an individual human to suffer from a power outage is proportional to a product of two factors: The overall outage probability, and the probability that the outage affects that specific area. This all suggests that the assumed linearity of the cybernetic models seems to collapse. Of course, multiplicative models are still linearizable and linear models are locally applicable, but here the goal is to derive global models over the whole operating regime; some more analysis is needed.

To continue with the cybernetic considerations, for a moment forget about the linear model structure, but concentrate on the essence, that is, on *balances.* Assume that in the cybernetic system the relevances are in static balance, that is, a set of equations holds:

$$\left\{ \begin{array}{c} \left(\frac{z_1}{\bar{z}_1}\right)^{a_{11}} \cdots \left(\frac{z_n}{\bar{z}_n}\right)^{a_{1n}} = \alpha_1 \left(\frac{\mu_1}{\bar{\mu}_1}\right)^{b_{11}} \cdots \left(\frac{\mu_m}{\bar{\mu}_m}\right)^{b_{1m}} \\ \vdots \\ \left(\frac{z_1}{\bar{z}_1}\right)^{a_{n1}} \cdots \left(\frac{z_n}{\bar{z}_n}\right)^{a_{nn}} = \alpha_n \left(\frac{\mu_1}{\bar{\mu}_1}\right)^{b_{n1}} \cdots \left(\frac{\mu_m}{\bar{\mu}_m}\right)^{b_{nm}} . \end{array} \right. \tag{7}$$

Here, $z_i$ are represent (positive) system variables and $\mu_j$ are external inputs; $\bar{z}_i$ represent the nominal values of the system variables. Taking logarithms one has

$$\left\{ \begin{array}{c} a_{11} \log z_1 + \cdots + a_{1n} \log z_n = c_1 + b_{11} \log \mu_1 + \cdots + b_{1m} \log \mu_m \\ \vdots \\ a_{n1} \log z_1 + \cdots + a_{nn} \log z_n = c_n + b_{n1} \log \mu_1 + \cdots + b_{nm} \log \mu_m, \end{array} \right. \tag{8}$$

where the constants $c_i$ contain the contributions of the operating point values $\log \bar{z}_i$ and $\log \bar{\mu}_j$ and other constants as collected together. This linear set of equations can be written in a matrix form

$$Ax = Bu, \tag{9}$$

where the constant terms have been combined in the variables, so that $x = \log z - A^{-1}c$, assuming invertibility of $A$. The steady-state of the vector $x$ can be solved as $\bar{x} = A^{-1}B\,u$. This seems familiar; it is evident that (assuming that $A$ can be interpreted as a stable system matrix) the originally static framework can be changed into a dynamic equilibrium process striving towards a cybernetic balance:

$$\frac{d\,x}{dt}(t) = -Ax(t) + Bu. \tag{10}$$

The linear cybernetic system structure has also been recovered. Again, interesting possibilities for studying networks in terms of "higher-order probabilistic balances" are available following the ideas in [6]. The main difference as compared to the populations-oriented models is that now all variables are *logarithmic,* so that, for example, $u_j = \log \mu_j$.

The multiplicative nature of systems, and the logarithmic nature of variables, is natural when studying proportions or probabilities, or relevances, rather than actual quantities or populations. It turns out that also in chemical equilibrium systems this kind of modeling is appropriate (see [12]). What is more, in many cybernetic systems involving humans (for example, in different kinds of decision networks) the logarithmic nature of variables are appropriate: It seems that humans naturally perceive things in logarithmic scales. This has been proven, for example, when studying concrete visual intensity sensitivity, but it can be claimed that our blindness to large quantities ("number dumbness" what comes to assessing numeric quantities beyond everyday scales) is related to the logarithmic nature of cognitive processing in general.

The model structure also remains intact, no matter if the variables are linear or logarithmic, and models can be constructed in the same way. However, cybernetic subsystems with different types of variables are mutually incompatible — they cannot be directly connected without variable transformations.

### 2.5 Properties of distributions

The simple structure between $u$ and $x$ makes it easy, for example, to draw conclusions concerning the statistical properties of the variable distributions. The key point is the linearity of the mapping.

Essentially, $\bar{x}$ is a weighted sum of (more or less) independent variables $u_i$ that are (more or less) equally distributed. If there are dozens of such input variables, it can be assumed (according to the *central limit theorem*) that the distribution of the sum approximates Gaussian, or normal distribution, so that the cybernetic variable $x_i$ has the distribution of the form:

$$p(x_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{(x_i - \mathrm{E}\{x_i\})^2}{\sigma^2}\right). \tag{11}$$

However, it has been recognized that complex variables typically do not have normal distribution; it seems that the "tails" are emphasized, so that there typically are "too many" big events. For example, in Internet, in another type of cybernetic networks, some nodes are visited extremely often, whereas there are very few hits for most of the nodes.

If the cybernetic variables are not normal, there must be something wrong with the assumptions? Not necessarily. Study a cybernetic system with logarithmic variables, so that $x_i = \log z_i$; the distribution is then *lognormal.* Take logarithms of (11), giving

$$\log p(\log z_i) = \alpha - \beta \left(\log z_i - \mathrm{E}\{\log z_i\})^2\right) . \tag{12}$$

It seems that if drawn on the log-log scale, the dependency is *quadratic.* This is interesting — rather than being linear (as a complex variable representing a self-similar fractal phenomenon should be) the dependency is parabolic on the log-log scale. There are three free parameters to match the curve against observations.

For example, if $z_i$ is the probability for an individual to be affected in a power outage, being proportional to the extent of the outage, the curve does not need to be linear — and, indeed, looking at Fig. 3, it is evident that the quadratic formulation is better justified. What is more, it seems that all of the above celebrated examples of scale-free nature (Figs. 1 and 2), can better be matched against parabolas than against straight lines! This is a relief really: If one wants to do quantitative analysis, one does not necessarily have to stick to the scale-freeness assumption. The framework of cybernetic systems offers an alternative route to analysis.

## 3 Synthesis of not-yet-existing systems

The cybernetic intuitions can be utilized also in design of systems. The above discussions were very descriptive, studying a large-scale system in a rather qualitative way; for more concrete discussions, quantitative studies are necessary.

### 3.1 Modeling distributed energy production

The idea of cybernetic higher-ordes balances offers a promising way to attack the problems of agent systems and distributed networks. The energy network itself is a cybernetic system, (hopefully) searching its balance in differing environments. To evaluate the claim that cybernetic considerations can be of some practical use, a concrete power plant simulation case was implemented.

It is assumed that there are $n$ power producers and $m$ power consumers, so that $n < m$. The productions are represented by the variables $\bar{x}_i$, where $1 \leq i \leq n$, and the consumptions are represented by the variables $u_j$, where $1 \leq j \leq m$. Consumers each have more or less random consumption realizations; however, there are periodic fluctuations that are characteristic to each consumer, but there are also correlations among different consumers. The demand has to be balanced at each time instant. Because the productions and consumptions can be directly summed together, linear variables are here appropriate.

Different kinds of production strategies were experimented. First, a centralized strategy was implemented: The system was explicitly optimized applying the known production costs at each individual energy producer. Second, a decentralized strategy was implemented, so that different producers were allocated mainly for specific consumers only. It is usually assumed that decentralization results in added roustness — but this holds only if the decentralization is implemented in a smart way. It is this smart distribution where cybernetic studies can be applied.

Mathematics helps to define the boundary line between the "easy" and the "difficult" problems. Difficult ones are those life-like problems that cannot be formulated explicitly; whenever the problem can be pinpointed, there are methods for optimizing behaviors. For example, robust control problems are solved routinely; the main difficulty is to formulate what *robustness* is in the first place. Fault tolerance in a system is an emergent phenomenon, being a result of the structural designs and parameter values, still missing explicit formulae.

It is a hunch that cybernetically designed "life-like" systems are *somehow* better than non-cybernetic ones; this intuition is evaluated in what follows.

### 3.2 Cost criteria

Typically, technical tasks can be formulated as optimization problems. When optimality is formulated, the system performance can be enhanced, either by explicit optimization methods, or by extensive number crunching. The key question is how to formulate the optimality criteria appropriately.

When formulating the energy production problem, the straightforward strategy is to write down the production costs at each production plant. If it is assumed that there is no cost for energy transfer, and there is no explicit cost for running up or shutting down of a plant, the cost criterion can be expressed in a static additive form as

$$C_{\mathrm{O}}(x) = \sum_{i=1}^{n} C_i(x). \tag{13}$$

On the other hand, if the system is explicitly decentralized, one can define fixed *production profiles* for each producer $i$ in the vector form $\varphi_i$. The elements in this $m$ dimensional vector reveal how relevant a specific consumer's needs are when the production is determined. Individual profiles can be combined in a matrix $\varphi$. To utilize such profiles in optimization, one can formulate a quadratic criterion that tries to compose the overall energy needs applying the available production profiles:

$$C_{\mathrm{D}}(x) = (u - \varphi x)^T \, (u - \varphi x) \,. \tag{14}$$

Cybernetically proper systems (according to [6]) are distributed systems where, again, the cost criterion has the same quadratic basic form

$$C_{\mathrm{C}}(x) = (u - \phi x)^T \, W \, (u - \phi x) \,. \tag{15}$$

To reach the "second-order balance", one has to select the matrix of profiles as follows

$$\phi^T = \mathrm{E}\{\bar{x}\bar{x}^T\}^{-1}\mathrm{E}\{\bar{x}u^T\}, \tag{16}$$

where $\bar{x}$ represents the steady-state values, and the weighting matrix is

$$W = \mathrm{E}\{uu^T\}. \tag{17}$$

The presented optimality criteria are mutually inconsistent. To achieve a well-defined optimization problem, the different criteria have to be combined, for example, by appropriate weighting, so that the final cost criterion becomes:

$$\begin{aligned} J(x,\lambda) = {} & C_{\mathrm{O}}(x) && \text{Production cost} \\ & + \rho\,(u-\varphi x)^T\,(u-\varphi x) && \text{Predetermined profiles} \\ & + \sigma\,(u-\phi x)^T\,W\,(u-\phi x) && \text{Cybernetic cost.} \end{aligned} \tag{18}$$

Here $\rho$ and $\sigma$ are weighting factors for emphasizing different criteria. Because there are constraints (for example, there must hold $\sum_i \bar{x}_i = \sum_j u_j$), one can introduce the set of constraints in the linear form

$$Gx = g. \tag{19}$$

Because of the extra terms in the cost criterion, seen from the perspective of production optimality, the system minimizing the combined criterion is suboptimal, and because of the non-cybernetic criteria, the system is also "sub-cybernetic", not truly reaching the balance goal. Following the discussions in [6], the external optimality criteria determine the "deprivation function" for cybernetic agents.

### 3.3 Minimizing the criterion

When the constrained optimization criterion is written as a single criterion applying the Lagrangian technique, one has

$$\begin{aligned} J(x,\lambda) = {} & C_{\mathrm{O}}(x) \\ & + \rho\,(u-\varphi x)^T\,(u-\varphi x) \\ & + \sigma\,(u-\phi x)^T\,W\,(u-\phi x) \\ & + \lambda^T(Gx-g). \end{aligned} \tag{20}$$

The criterion can be differentiated with respect to the variables:

$$\begin{cases} \dfrac{dJ}{dx}(x,\lambda) = \frac{dC_{\mathrm{O}}}{dx}(x) - 2\rho\varphi^T(u-\varphi x) - 2\sigma\phi^T(u-\phi x) + G^T\lambda \\ \dfrac{dJ}{d\lambda}(x,\lambda) = Gx - g. \end{cases} \tag{21}$$

In principle, the criterion can be minimized applying the gradient descent algorithm (expressed in continuous time below), so that one lets the process converge to its fixed state:

$$\begin{cases} \dfrac{dx}{dt}(t) = -\gamma\frac{dJ}{dx}(x,\lambda) \\ \qquad\quad = -\gamma\left(\frac{dC_{\mathrm{O}}}{dx}(x) - 2\rho\varphi^T\,(u-\varphi x) - 2\sigma\phi^T\,(u-\phi x) + G^T\lambda\right) \\ \dfrac{d\lambda}{dt}(t) = -\gamma\frac{dJ}{d\lambda}(x,\lambda) \\ \qquad\quad = -\gamma\,(Gx-g)\,. \end{cases} \tag{22}$$

However, the dynamics of the above process can be complicated; furthermore, there typically exist various local minima for complex $C_{\mathrm{O}}(x)$, and typically it is not continuously differentiable. It is necessary to elaborate on the optimization a bit more. The optimum is either in some point where there does not exist gradient, or then there holds

$$\frac{d\,J}{d\,x}(\bar{x},\bar{\lambda}) = 0. \tag{23}$$

Studying (22) gives

$$\bar{x} = \left(2\rho\varphi^T\varphi + 2\sigma\phi^T W\phi\right)^{-1}\left(2\rho\varphi^T u + 2\sigma\phi^T W u - G^T\bar{\lambda} - \frac{d\,C_{\mathrm{O}}}{d\,x}(\bar{x})\right). \tag{24}$$

Taking $G\bar{x} = g$ into account, one has

$$\begin{aligned}\bar{\lambda} = &\left(G\left(2\rho\varphi^T\varphi + 2\sigma\phi^T W\phi\right)^{-1} G^T\right)^{-1} \\ &\quad\left(G\left(2\rho\varphi^T\varphi + 2\sigma\phi^T W\phi\right)^{-1}\right. \\ &\qquad\left.\left(2\rho\varphi^T u + 2\sigma\phi^T W u - \tfrac{d\,C_{\mathrm{O}}}{d\,x}(\bar{x})\right) - g\right).\end{aligned} \tag{25}$$

For a given consumption pattern $u$ one first calculates $\bar{\lambda}$ from (25), and then $\bar{x}$ from (24).

A simple one-step procedure for determining the minimum can be carried out only in simple cases. The main problems that remain are caused by the nature of the optimality criterion $C_{\mathrm{O}}$. A rather plausible model is the piecewise affine cost criterion

$$C_{\mathrm{O}}(x) = \sum_{i=1}^{n} \delta_i(x_i)\,(a_i + b_i x_i). \tag{26}$$

This means that for each power plant $i$ there is a constant cost $a_i$ for keeping it running, and the cost for each produced unit of energy is determined by the parameter $b_i$. For simplicity, the maximum cannot be here exceeded; in practice, the cost for extra production would only start rising steeply. This could be analyzed by introducing yet other locally linear regions in the cost model. It is assumed that if there is no production, there is no cost; for this reason, the additional factor $\delta_i$ is included in the model:

$$\delta(x_i) = \begin{cases} 1, & \text{if } x_{\mathtt{min},i} \leq x_i \leq x_{\mathrm{max},i}, \\ \infty, & \text{if } x_i \geq x_{\mathrm{max},i}, \text{ and} \\ 0, & \text{if } x_i = 0. \end{cases} \tag{27}$$

The gradient of the presented criterion in the active area (where all of the producers $i$ remain between $x_{\mathrm{min},i}$ and $x_{\mathrm{max},i}$) is a constant vector

$$\frac{d\,C_{\mathrm{O}}}{d\,x}(x) = b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}. \tag{28}$$

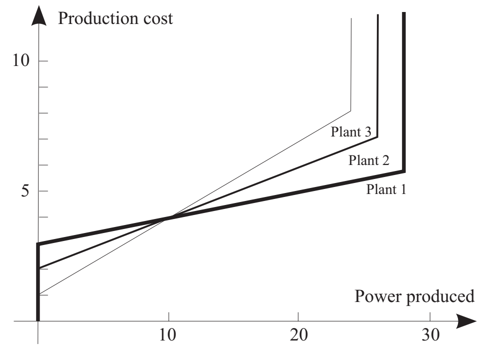


**Fig. 5.** The cost profiles for the three producers

Because of the piecewise linear nature of the cost criterion, minimization has to be repeated for each basin of local minima separately: All combinations of each of the variables $x_i$ being either minimum, maximum, or in active state, have to be studied, and (25) and (24) are calculated. From these, the global minimum is selected, calculating (20) for each valid candidate $\bar{x}$. A candidate is valid only if the assumptions hold: The solution has to remain between the operational range of individual power plants, that is, there must hold $x_{\min,i} \leq \bar{x}_i \leq x_{\max,i}$ for all $i$.

The different operating regimes can easiest be implemented by integrating the modes of operation in the constraint matrices $G$ and $g$. These data structures can be constructed iteratively as follows, starting from empty matrices:

> Let the "constraint index" be $c = 0$. As long as there are unprocessed constraints, let $c = c + 1$, and modify data structures as follows:
> 1. If the variable $x_i$ is fixed in minimum, let $G_{ji} = 1$ and $g_j = x_{\min,i}$.
> 2. If the variable $x_i$ is fixed in maximum, let $G_{ji} = 1$ and $g_j = x_{\max,i}$.
> 3. Finally, implement the production constraint: Let $G_{ji} = 1$ for all $i$, and let $g_j = \sum_{j=1}^{m} u_j$.

### 3.4 Simulations and observations

In the simulations there were $n = 3$ producers and $m = 20$ consumers. In Fig. 6, a typical realization of the consumer behaviors is shown; clearly, there are cyclic patterns, and there are redundancies among the behaviors.

The simulation results are shown in Figs. 7, 8, and 9; the consumption in each case is the same, shown in Fig. 6, and at each time instant the total production equals the total consumption. In Fig. 7 one has $\rho = \sigma = 0$, so that the strategy is strictly optimized — the cost for producing energy in each plant is shown in Fig. 5. It turns out out that there are violent variations in the production patterns when this strategy is applied; only one producer is active at a time, the other ones being in either extremum. The explicitly distributed strategy is shown in Fig. 8, meaning that $\sigma = 0$ whereas $\rho$ is large, applying a more or less random distribution strategy. It turns out that the production style is far from optimum, and, again, there are violent variations.

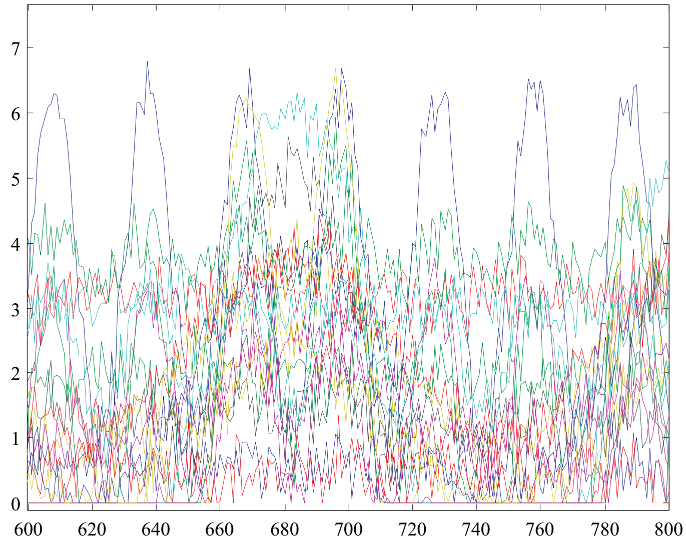


**Fig. 6.** Behaviors of 20 energy consumers

Finally, the cybernetic case was simulated by setting $\rho = 0$, and letting $\sigma$ change from 0 to a large value. The reason for this is that because the system is nonlinear, there are various alternative operation regimes, determination of the correlation matrices must be *constructivistic,* dependent of the history. After a few iterations the profiles $\phi$ converge to match the statistical properties of the data. The result is shown in Fig. 9: It turns out that the production style is near optimal, but the variations in the production patterns are much smoother than in the strictly optimal case. Indeed, the production style resembles real-life behaviors as actual production systems are observed.

The simulations give some intuition of what is the nature of *robustness* in life-like networks (assuming that the adopted view of how cybernetics should be characterized and modeled is correct):

- Globally optimized control always runs the power plants so that only one of them is active at a time, others being in their extreme values (assuming affine cost increase between minimum and maximum). The cybernetic scheme, on the other hand, seems to avoid extreme values. It is evident that as more plants are active, the more there is buffer against sudden changes in consumption.
- Because the profiles are based on (sparse) principal components, the plants are insensitive against random noise (compare to principal component analysis). The plants only react to *real* underlying changes in consumption, probably resulting in smoother production.
- Explicitly distributed systems with some fixed profiles (for example, a few plants taking care of a set of consumers) is vulnerable to domino effects: If one of the plants is out, the sudden excessive load can collapse the other ones,

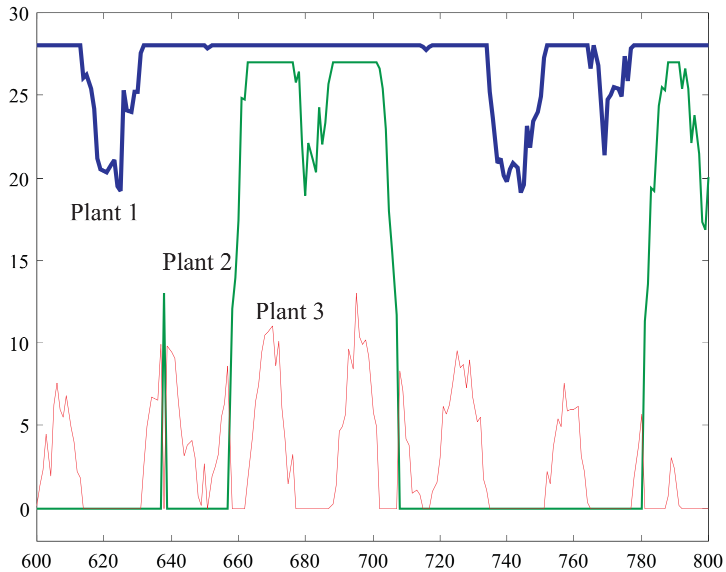


**Fig. 7.** Strictly optimized production

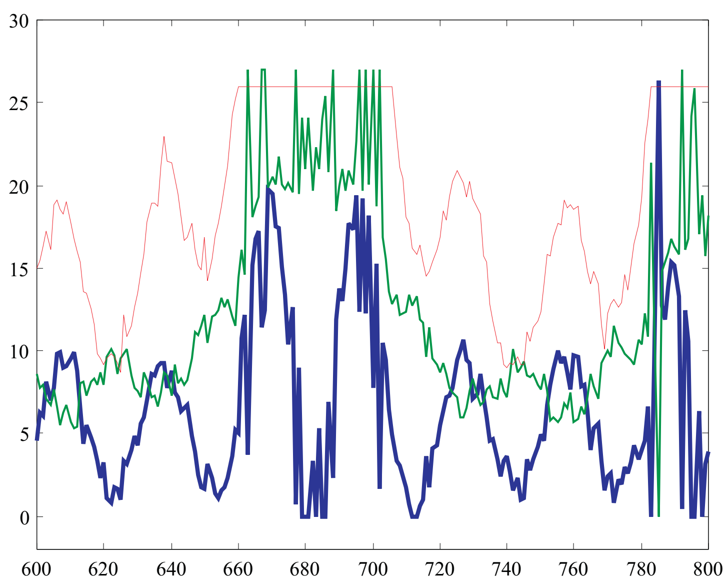


**Fig. 8.** Distributed production

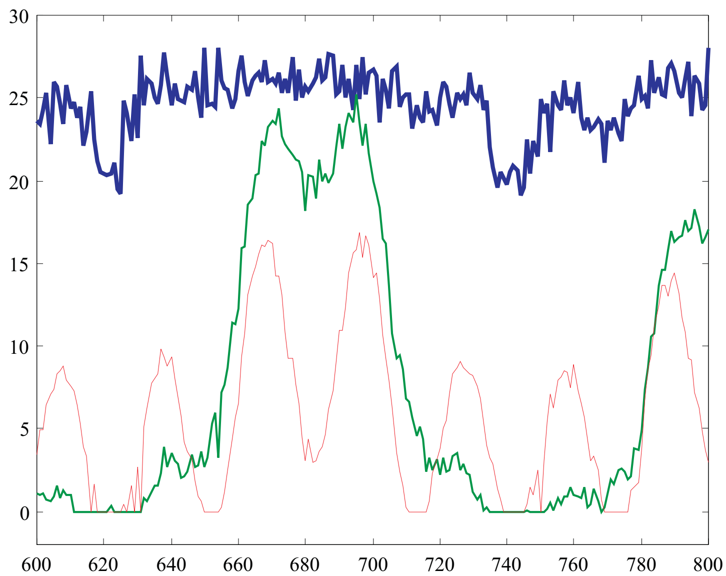


**Fig. 9.** "Cybernetic" production

too. Now, the profiles are based on (sparse) principal components, meaning that the profiles are (almost) orthogonal. Collapse of one plant does not excessively strain any of the other plants alone; rather, the additional load is distributed evenly among the reserve plants[1].

### 3.5 Design of networks

What technical use there is for cybernetic models? In the vicinity of the absolute optimum, there are still degrees of design freedom available: System optimality not too severely compromized if minor modifications are done, and different kinds of enhancements can be proposed. Assuming that the system robustness is related to the probability of sudden changes in the system variables — rapid variations should be minimized —, cybernetic models directly seem to give tools for optimizing robustness. Different kinds of sensitivity analyses are possible, and when the system is structured in a new way, one can find the most relevant parameters affecting the network behavior.

Above, it was assumed that there are no transfer costs. In this sense, the models are information theoretic, not bound to physical constraints, locations or distances. Inversely, if the network is just being designed, cybernetical models can be applied for determining the plant locations: When the environment is

[1] Because of unoptimal distribution of load, the level of random variation increases in all plant activities; indeed, similar fluctuations are detected in ecosystems if some key species disappears

analyzed cybernetically, and the nodes are found, the best locations for them can be decided based on physical criteria.

There are very different types of networks. For example, as compared to energy production networks, in Internet one is actually facing an inverse situation: There exist no practical limitations for production (copying of files) but the data transfer rates are limited. Rather than modeling the nodes, one should model the arcs in the network. In this sense, one can draw a *dual network* where the nodes are arcs, and *vice versa,* and apply the same modeling principles as above — not for determining the production profiles, but for determining the distribution of load within the transfer channels.

Networks can be analyzed applying traditional optimization techniques if the cost criterion is defined; the problem in the presented case was that of determining the criterion. However, in some networks there does not seem to exist meaningful global criteria, and decentralized strategies only are applicable. Another example of cybernetic networks that is studied in [12] is a distributed sensor network. If the sensors are fully connected, the network carries out principal component filtering; this behavior is trivial, being the same as when using a centralized architecture. More interesting functionalities emerge if the network of sensors is *not* fully connected. If only the nearest neighbors are connected, the latent variables become localized in an interesting way, and, even though information is incomplete, the resulting estimates seem to be *more* accurate than in the centralized implementation. The cybernetic framework seems to give new substance to the agent paradigm.

## References


1. Johnson, S., *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*, Touchstone, 2002.
2. Barabasi, A-L., *Linked: The New Science of Networks*, Perseus Publishing, 2002.
3. Bunde, A. and Havlin, S. (Eds.): *Fractals in Science.* Springer Verlag, 1994.
4. Carlson, J.M. and Doyle, J.: *Highly Optimized Tolerance: Robustness and Design in Complex Systems.* 1999.
5. Carreras, B.A., Newman, D.E., Dobson, I., and Pooloe, A.B.: Evidence for Self-Organized Criticality in a Time Series of Electric Power System Blackouts. To appera in *IEEE Transactions on Circuits and Systems,* 2004.
6. Hyötyniemi, H.: Cybernetics — Towards a Unified Model? Submitted to the *Finnish Artificial Intelligence Conference (STeP'04),* Vantaa, Finland (September 2004).
7. Malamud, B.D., Morein, G., and Turcotte D.L.: *Forest Fires: An Example of Self-Organized Critical Behavior. Science,* Vol. 281, Issue 5384, pp. 1840–1842, September 18, 1998.
8. Pearl, J: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1997.
9. Saaty, T.L. and Vargas, L.G.: *Models, Methods, Concepts & Application of the Analytic Hierarchy Process.* Kluwer Academic Publishers Boston, 2001.
10. Senge, P.M.: *The Fifth Discipline: The Art and Practice of the Learning Organization.* Doubleday Currency, New York, 1990.

11. Simon, H.A.: *Sciences of the Artificial.* MIT Press, 1969 (first edition).

12. *Additional material on cybernetics will be available in public domain in near future at* `http://www.control.hut.fi/cybernetics`*.*