Parameter-tuning networks: Experiments and active-walk model

XIAO-PU HAN\textsuperscript{(a)}, CHUN-DONG Hu\textsuperscript{2}, ZHI-MIN LIU\textsuperscript{2} and BING-HONG WANG\textsuperscript{1,3}

\textsuperscript{1} Department of Modern Physics, University of Science and Technology of China - Hefei, 230026, China
\textsuperscript{2} Institute of Plasma Physics, Chinese Academy of Sciences - Hefei, 230031, China
\textsuperscript{3} Shanghai Academy of System Science and University of Shanghai for Science and Technology Shanghai, 230093, China

received 23 March 2008; accepted in final form 2 June 2008
published online 2 July 2008

PACS 89.75.-k – Complex systems
PACS 89.75.Hc – Networks and genealogical trees
PACS 05.40.Fb – Random walks and Levy flights

Abstract – The tuning process of a large apparatus of many components could be represented and quantified by constructing parameter-tuning networks. The experimental tuning of the ion source of the neutral beam injector of the HT-7 Tokamak is presented as an example. Stretched-exponential cumulative degree distributions are found in the parameter-tuning networks. An active-walk model with eight walkers is constructed. Each active walker is a particle moving with friction in an energy landscape; the landscape is modified by the collective action of all the walkers. Numerical simulations show that the parameter-tuning networks generated by the model also give stretched exponential functions, in good agreement with experiments. Our methods provide a new insight to understand the action of humans in the parameter-tuning of experimental processes, and our model could be helpful for the experimental research and other optimization problems.

Copyright © EPLA, 2008

Introduction. – Large apparatuses play more and more important roles in the experimental research of modern physics. For example, the advancement of high-energy accelerators and of the large devices for controlling nuclear fusion is the necessary impetus for the research of particle physics and plasma physics. In a large apparatus made up of many components, each component needs be tested and tuned separately and then collectively, before experiments, using the apparatus as a whole, can be conducted. And there are a large number of parameters to be tuned. The parameter tuning processes are often complicated and onerous, and strongly depend on the experience of the experimenter. How can we model the adjustment process of these parameters and make some sense out of it?

For example, in a Tokamak in plasma physics, in the ion source segment of the neutral-beam injector system alone, there are eight major parameters to be adjusted. The experimenter sets the parameter values, turns on the equipment, and measures the outcome of a certain quantity $Q$. If the $Q$ obtained does not meet the mark $Q_0$, say, the whole process is repeated, with a new set of parameters. The adjustment process stops when $Q$ is equal or very close to $Q_0$. The simultaneous adjustment of a large number of parameters is a complicated process, which is based on the feedback from previous $Q$ values obtained, the experimenter’s experience, and the limitation of the hardware that control the parameters.

Parameter-tuning networks. – To quantify this complicated process, a parameter-tuning network can be constructed as follows. Assume that $N$ parameters $u_i (i = 1, 2, \ldots, N)$ are involved. The experimenter’s serial adjustment of $u \equiv (u_1, \ldots, u_N)$ can be represented by a sequence of connected dots in the $N$-dimensional $u$ space. For each choice of $u$, $Q$ is measured, so that $Q = Q(u)$ but the functional form is unknown due to the complexity of the equipment, which is like a black box.

The sequence of $u$ dots is then projected onto the $(u_1, u_2)$-plane, say. The dots on the same vertical line along the $u_3$-direction are allowed to collapse to one point, called a node. A line connecting two nodes is called an edge [1–3]. Since in real situations, the parameters selected by the experimenter in different adjustments may partially overlap with each other, there may be more than one edge connecting two nodes. These edges are directed. For simplicity, we make the approximation of collapsing all the edges between two nodes that have the same direction as one single edge with the same direction. Consequently, between two nodes, there are
at most two edges with opposite directions, resulting in a directed parameter-tuning network. A non-directed network is formed from the directed one, by collapsing all the edges between any two nodes and removing the directions.

In a parameter-tuning network, there are \( N_n \) nodes, say. For a node \( j (j = 1, 2, \ldots, N_n) \) in the directed parameter-tuning network, let \( N_+(j) \) and \( N_-(j) \), respectively, denote the number of ingoing edges and outgoing edges.

Now count the number of \( N_+ \) that has the value \( k \), and call it \( p_+(k) \); similarly for \( p_-(k) \). The corresponding function in the non-directed network case is denoted by \( p(k) \). \( k \) is called the “degree” of a node. \( p_+(k) \) is called the “in-degree distribution”; \( p_-(k) \) the “out-degree distribution”; and \( p(k) \) the “degree distribution.” Next define the corresponding “cumulative degree distributions” \( P_+(k) \), \( P_-(k) \) and \( P(k) \) by

\[
P_+(k) \equiv \sum_{k' = k}^{k_m} p_+(k'),
\]

\[
P_-(k) \equiv \sum_{k' = k}^{k_m} p_-(k'),
\]

\[
P(k) \equiv \sum_{k' = k}^{k_m} p(k'),
\]

Here \( k_m \) is the maximum \( k \) at which the \( p \) function is non-zero, and all of the distribution functions are not unitary. Note that by definition, the \( P \)’s are monotonic decreasing functions, while the \( p \)’s may not be monotonic at all. The reason for introducing the \( P \)’s is that we want to fit them to monotonic decreasing functions such as stretched exponents or power laws.

In our experiments studying the ion source of the neutral beam injector system for the HT-7 Tokamak, eight control parameters are involved \( (N = 8) \). These include the filament current, magnet current, arc voltage, cathode gas valve voltage, anode gas valve voltage, etc. Each of the parameters has 10 to 30 discrete, adjustable setting values. The aim of the experiment is to find out which set of parameters will give a strong and stable discharge, measured by the arc current intensity \( Q \). Each parameter is set by a dial which can be turned left or right between two extreme positions. When the extreme position is reached, the experimenter has to turn the dial back and reverses the direction of turning.

From our experimental data, the sequence of parameters in the 8-dimensional \( u \) space is first generated and the corresponding directed and non-directed parameter-tuning networks are constructed. The non-directed network is shown in fig. 1. The \( P \)’s are obtained, and fitted nicely with stretched exponential functions (fig. 2).
such that
\[ P(k) \sim \exp(-ak^\gamma) \] (4)
or, equivalently,
\[ \ln P(k) = -ak^\gamma + b \] (5)
and similarly for \( P_+(k) \) and \( P_-(k) \). Such property is also found in the parameter-tuning networks derived from the data of the tuning experiments in other experiment season of the ion source (each of them includes 600 to 800 parameter sets).

The model. – The experimental adjustment of the parameters is somehow correlated with the experimenter, and the process can be modeled by an active-walk (AW) model [4,5]. In our AW model here, \( N = 8 \); each \( u_i \) has adjustable values of \( 1, 2, \ldots, 20 \), say. The adjustment of the \( i \)-th parameter is represented by the movement of an active walker on a \( 1D \) landscape potential \( V_i(x_i) \). The allowable \( x_i \) for the \( i \)-th walker are the integers \( 1, 2, 3, \ldots, 100 \), the same for all \( i \)'s. These 100 numbers are mapped to the \( u_i \), such that if the \( i \)-th walker ends in the region \([1 \sim 5] \) on the \( x_i \)-axis, \( u_i \) will assume the value 1. Similarly, the region of \([6 \sim 10] \) is mapped to \( u_i = 2 \), etc. This mapping can be attained by
\[ u_i(t) = \text{Int}[(X_i(t) - 1)/5] + 1, \] (6)
where \( X_i(t) \) is the position of the \( i \)-th walker on the \( x_i \) axis at time \( t \), and \( \text{Int} \) means taking the integral part of the number. This mapping of \( x_i \) to \( u_i \) has the effect of making two consecutive sets of adjusted parameters more likely to partially overlap with each other, and ensures the occurrence of smooth \( V_i \) vs. \( x_i \) curves.

In the simulations, all \( V_i 's \) start flat at time \( t = 0 \), and are updated simultaneously at each time step \( t (=1, 2, \ldots) \), according to a rule to be specified below. But between two consecutive time steps, each walker moves a few steps in a subwalk in its own \( V_i \), independent of other walkers. In a subwalk, the walker does not change \( V_i \). The subwalk is like a particle rolling on a landscape with friction, with the following rules (with the subscript \( i \) removed for the sake of clarity). The subwalk time is labeled by \( \tau \) (= 0, 1, 2, \ldots).

i) At time \( t = 0 \), the particle is arbitrarily placed on the \( x \)-axis.

ii) At time \( t \), the particle is given energy \( K_0 \) at \( \tau = 0 \). (\( K_0 \) is a parameter fixed in the model.)

iii) At subwalk time \( \tau \), the particle moves left or right with equal probability. However, after each move, the particle loses energy \( \varepsilon \), and the energy difference \( V(x(\tau)) - V(x(\tau + 1)) \) which could be positive or negative is added to its energy. Consequently, at time \( \tau \), the particle already moves \( \tau \) steps, and its energy becomes
\[ K(\tau) = K_0 + V_0 - V(x(\tau)) - \tau \varepsilon, \] (7)

where \( V_0 \equiv V(x(0)) \), the potential at the initial position of the particle at \( \tau = 0 \).

iv) The particle can get over a potential barrier in \( V \) that is lower than \( K \), and can rebound if the barrier is higher than \( K \).

v) When the particle reaches the left boundary \( (x = 1) \) or the right boundary \( (x = 100) \), it reverses direction and continues its subwalk.

vi) The particle stops when it walks into a potential well and cannot get out with its energy \( K \), or exhausts its energy on a plateau.

These possibilities are sketched in fig. 3. The reflecting boundary condition in item v) corresponds to the real experimental situation where a dial with a limited range is used.

After all \( N \) particles stop, the time clock increases from \( t \) to \( t + 1 \); the particle's position at the end of its subwalk is taken to be \( X(t + 1) \), which is the starting position of the subwalk at time \( t + 1 \). (The subwalks of the particles may not stop after the same number of subwalk steps; those stop first will sit there and wait for the last particle to stop.) The landscape \( V_i(x_i) \) is updated by the landscaping rule:
\[ V_i(x_i; t + 1) = \begin{cases} V_i(x_i; t) + W(x_i - X_i(t + 1)), & E(t) \geq E_0(t), \\ V_i(x_i; t) - W(x_i - X_i(t + 1)), & E(t) < E_0(t), \end{cases} \] (8)
where \( t = 0, 1, 2, \ldots \). In our numerical simulations below, the \( (i \)-independent) landscaping function \( W \) is given by \( W(0) = 4, W(\pm 1) = 3, W(\pm 2) = 1 \), and \( W = 0 \), otherwise. In eq. (8), the error function \( E(t) \) is assumed to be
\[ E(t) = |Q(u(t)) - Q_0|, \] (9)
where \( u(t) \) is obtained from eq. (6); \( E_a(t) \) is the average error counting the last \( n \) time steps, given by

\[
E_a(t) = \frac{1}{n} \sum_{t' = t-n+1}^{t} E(t').
\] (10)

In computer simulations, we pretend that we know \( Q_0 \) and the \( Q(u) \) function. In reality, the former is known to the experimenter; the latter can be roughly inferred from experimental data. Note that while the subwalks of the particles are independent of each other, the updating of their landscapes is affected by their collective effort through \( Q(u) \) in \( E(t) \).

The eight-parameter AW model is simulated with \( t \) going from 0, and the simulation is stopped when the optimal parameter set is gotten. The function \( Q \) is assumed to be

\[
Q(u) = |u_1 - 10| + |u_2 - 10| + \cdots + |u_8 - 10|.
\] (11)

The parameters used are: \( K_0 = 23, \varepsilon = 1, n = 5 \) and \( Q_0 = 0 \). This means that \( u = (10, 10, \ldots, 10) \) is the one and only optimal parameter set.

The error function \( E(t) \) is displayed in fig. 4, which shows that the optimal \( u \) is first obtained at \( t_0 = 3718 \). The 700 \( u(t) \) dots (with \( 300 \leq t < 1000 \), this number of \( u(t) \) is similar with the experimental data, and the initialization process when \( t < 300 \) is ignored) in the \( u \) space are projected onto the \((u_1, u_2)\)-plane to obtain the directed and non-directed networks. The corresponding cumulative degree distributions are given in fig. 5. Good stretched-exponential fits are obtained, in agreement with the experimental findings in fig. 2.

How to understand the underlying mechanism of the stretched-exponential properties in the AW model and experimental results. Because of the subwalk process and the averaging effect in eq. (10), the model is non-Markovian. This makes the analytical study of the model quite difficult. But we can understand it from the evolution of the landscape \( V_i \). If the landscape is unchanged, the particles move randomly, and the model is retrogressed to a normal random walk, and the structure of the network is homogeneous. Actually, on the one hand, because of the changing landscape, after a long-term simulation, deep wells appear in the field that close to the optimal set, as shown in fig. 6. This implies the particle has more probability to stop at these wells. On the other hand, comparing with the initial states, the changing range of the potential of the positions with a deep well generally is larger than others. If a walker starts the next subwalk from a position, it only has at most two possible positions to stop (it is called stopping position in the following) since the initial energy is fixed. Because the landscape evolves step by step, if the position has a large changing range of potential, its cumulative number of the stopping positions generally is also large in the evolution process. According to the construction of the parameter-tuning network, the cumulative number of stopping positions is also positively related with the degree. These two reasons make the structure of the parameter-tuning network generated by the AW model inhomogeneous. In the real experimental tuning process, the experimenter usually tunes the parameter set more...
carefully and more minutely when the outcome is close to the aim. Thus the nodes (parameter sets) with a large degree usually have good outcomes. This phenomenon is also fitted by the AW model. In a word, the inhomogeneous structure of the parameter-tuning networks in both the AW model and experiments comes from the optimization of the system.

Conclusions. – In the AW model, the subwalks kind of mimic the action of the experimenter—we think that is why the AW model and the experiments give similar P functions. Human dynamics [6,7] have been studied in other situations by different models. Our results imply that our network-based statistics and the AW model is helpful in the understanding of human action in the experimental processes and very likely in many other optimization projects. The stretched exponent property of the parameter-tuning network imply the human action in the optimization process could be having some quantitative principles, which is important for the experimental studies. Stretched exponent distributions or relations exist widely in many other systems [8–11]. Since AW-like dynamics is very common in nature [4,5], our model may reveal a new mechanism of stretched exponent distribution, but the strict analytical research of the AW model needs further studies. By modifying the subwalks and tuning their parameters one may shorten $t_0$, the shortest time to find the optimal parameters, which would be of interest to the experimental researchers. Finally, the AW model mimics the process that optimize the parameter sets in real experiment, so it could be useful in the optimization projects of many other systems.

***

We thank LUI LAM, PIN-QUN JIANG, DA-REN HE and TAO ZHOU for useful discussions, and SHENG LIU, SHI-HUA SONG, JUN LI and YUAN-LAI XIE for experimental help, and JUN-FANG ZHU and LUO-LUO JIANG for other beneficial help. We also thank the anonymous referees for constructive criticisms and advices. This work is partially supported by China 973 Program (Grant No. 2006CB705500) and NSF China (Grant Nos. 60744003, 10635040, 10532060, 10472116, and 10575105).

REFERENCES

[1] WATTS D. J. and STROGTZ S. H., Nature, 393 (1998) 440.
[2] BARABÁSI A.-L. and ALBERT R., Science, 286 (1999) 509.
[3] ALBERT R. and BARABÁSI A.-L., Rev. Mod. Phys., 74 (2002) 47.
[4] LAM L., Int. J. Bifurcat. Chaos, 15 (2005) 2317.
[5] LAM L., Int. J. Bifurcat. Chaos, 16 (2006) 239.
[6] BARABÁSI A.-L., Nature, 435 (2005) 207.
[7] VÁZQUEZ A., Phys. Rev. Lett., 95 (2005) 248701.
[8] XULVI-BRUNET R. and SOKOLOV I. M., Phys. Rev. E, 66 (2002) 026118.
[9] HOLANDA A. J., PISA I. T., KINOUCHI O., MARTINEZ A. S. and RUIZ E. E. S., Physica A, 344 (2004) 530.
[10] LAHERRÉRE J. and SORNETTE D., Eur. Phys. J. B, 2 (1998) 525.
[11] STURMAN B., POVIDILOV E. and GORKUNOV M., Phys. Rev. Lett., 91 (2005) 176602.