UT-616
TMUP-HEL-9206
KEK-TH-341
KEK preprint 92-71

Power law in the human memory and in the neural network model[*] [‡]

Hideyuki CÂTEAU[a†]
Tatsuhiro NAKAJIMA[b]
Hiroshi NUNOKAWA[b,c†]
Nobuko FUCHIKAMI[b]

[a] *Department of Physics, University of Tokyo*
Bunkyo-ku,Hongo, Tokyo, 113 Japan
[b] *Department of Physics*
Tokyo Metropolitan University
1-1 Minami-Ohsawa, Hachioji,
Tokyo 192-03 Japan
[c] *National Laboratory for High energy physics (KEK)*
Tsukuba-shi, Ibaraki 305, Japan

July 1992

∗ The authors thank Dr.Love and Dr.Taki for their useful suggestions on the back propagation simulator. A part of the present work was done by using the neural network simulator Aspirin (R.Leighton, The MITRE Corporation). The authors also acknowledge valuable conversations with Professors A.Shinohara and S.Ichihara of the department of Psychology of Tokyo Metropolitan University. One of the authors (H.N) thanks M. Kobayashi for allowing him to visit KEK. This work was supported in part by *A Fund for Special Research Project at Tokyo Metropolitan University.*
‡ The reprint of this article is to be requested to H.Câteau, University of Tokyo,Hongo,Bunkyo-ku,Tokyo, 113 Japan,e-mail:cateau@star.phys.metro-u.ac.jp
† Fellow of Japan Society for the Promotion of Science.

Power law in the human memory and in the neural network model

Key words:back propagation, encoder, power law, psychology
memory model, scaling, physiology, learning curve

Abstract

We show that the learning pace of the back propagation model is described by a power law with high precision. Interestingly the same power law was found out in the human memory by a psychologist in the past. Therefore our result provides a quantitative evidence that the back propagation model, though it is simple, surely shares some essential structure with the human brain. In proceeding the discussion we propose a novel model of the memory system. The model overcomes the notable difficulty lying in the back propagation network that the learning time is very sensitive to the initial condition.

1. Introduction

The neural network model was originally proposed as a model of a brain of living things (McCulloch and Pitts 1943). However, the studies of it have been spread over many unexpected topics so far. Through such studies we have gained deep understanding on many aspects of the neural network model itself.

Here, let us consider what we should do next. In our opinion this is the time to address the original question again. Namely we ask whether the neural network model is really a good model of our brain or not. We ask whether the study of the neural network model casts a light on the secret of the human brain. The ample knowledge on the neural network model would help us tackle this ultimate question more seriously at present.

Qualitative studies indicating some affirmative answers to this question have been given so far. For example, it was pointed out (Rumelhart et al. 1986) that what is seen in the neural network model with its partial destruction resembles to what is seen in the human deep dyslexia. On the other hand no much attention has been paid to the quantitative aspects of the problem.

Instead of considering qualitative features, we focus our attention on a quantitative feature which is common between the neural network model and the human brain. Namely, we show in this paper that the memory performances of the both systems are subject to the same power law. The values of the exponents are also shown to be nearly equal between the two. By comparing the power laws of the human brain and the neural network model, we find the correspondence between the time scale of the human brain and the algorithm step of the neural network. For a large range of the parameter values of the model it is shown that our brain spends $o(1) \sim o(10)$ msec for a unit step learning. This reaction time of the neurons is physiologically natural (Feldman 1982,1985)(Lynch,Mountcastle,Talbot and Yin 1977). In proceeding the discussion we notice an important difference between a single back propagation (BP) network (Rumelhart, Hinton and Williams 1986, Widrow and Hoff 1960, Amari 1967) and the human brain. That is the stability of learning. As is well known, the learning time of BP is very sensitive to the choice of the initial condition, while the human learning time is not so unstable. We point out that a composite network of many BP networks fills up this gap in a natural way.

In the next section we introduce the power law in the human memory proposed by M.Foucault and present our reexamination of it. In the section three we show the same power law in the BP model. A novel memory model will be proposed therein. The last section is devoted to some discussions. A short account of this article is given in the refs. (Câteau et al. 1992, Nakajima et al. 1992)

2. Power law in the human memory

From our daily experience we know that when we memorize some items, the pace of the memory slows down as the number of items to memorize increases. Even if we can memorize 10 names of people in the time t_1, we cannot memorize 20 names within the time $2t_1$. We need longer time than $2t_1$. This phenomenon is understood in psychology to be a result of interference between different items to memorize. When we memorize a new item in addition to items which have been already memorized, we have to take care not to mix the new one with the old ones. This costs time and slows down the memory. The psychologist, M.Foucault expressed this fact as the following empirical law (Foucault 1913):

$$t(M) = cM^D \tag{2.1}$$

with $D = 2$, where M denotes the number of items to memorize, t denotes the time necessary to memorize them and c is a constant. Because Foucault's experiment is very old one, we first performed a reexamination of this law. We prepare a sheet of paper on which a sequence of random numbers is written. We ask a subject to memorize these numbers in order as much as possible. Every thirty seconds, we ask the subject to temporally stop the learning and we check how many numbers he/she has memorized.

Our psychological experiment was performed in this way. This provides the data of the form $\{(t_1, M_1), (t_2, M_2), \cdots\}$ for each subject. If the power law of eq. (2.1) holds, these data points, when plotted on a bilog graph, should exhibit a linear behavior. The slope of the plots corresponds to the exponent D. We accomplished this experiment for students in the Tokyo Metropolitan Unversiy, and plotted the data of them on a bilog graph. Displayed in Figure 1 are typical plots of such. Besides this we also present in Figure 2 the plots of the data from persons with abnormally high memory. Such data are found in the old psychological literature (M.Foucault 1913). Note the difference on the capacity of the memory between ordinary people and talents. In these figures, the error bars are obtained by assuming 10% error for each data point.

As you see, the linearity of the bilog plots seems good. To gie more qualitative justification, we performed a χ^2 test of these data by assuming two cases. One case employs the standard deviation obtained from 10% error assumed for each data point and the other corresponds to the assumption of 20% error. The significance level thus obtained is given in Table 1, which is mostly high except a few exceptions. In this way we would conclude that the linearity is fairly good, i.e. the power law dependence holds well.

If the power law woks universally among many people, that is rather amazing. Comparing the normal people with talents with abnormally high memory (see Figures 1 and 2), we may naturally believe that some essential difference must lie in functioning of their memory systems. Nevertheless, our experimental results indicate a common quantitative feature of the memory system which is expressed as the power law. In comparison between Figure 1 (normal people) and Figure 2 (talents), the values of c and D have no considerable difference. The great difference between the two is only the range of t and M. The talents can memorize many items spending the long time, while the normal people immediately reach their limit then the range of t and M are narrow.

Our experiment was extended over various types of students. Many aspects are different among these students such as age, sex, major of interest, physical or psychological conditions. As we usually feel, the power of the memory in general depends crucially on such factors. Furthermore the way to learn the random sequence is different for the respective subjects. Someone assigns some meaning to the random sequence to concrete his/her memory, while someone assigns merely a rhythm to it. It is clear that such tactics may improve the amount of the memory. Our result suggests that although the performance of the memory varies depending on a variety of individual states of the subjects or on such tactics, it is fairly universal that they are subject to the same type of power law expressed as equation (2.1).

Now we comment on the value of the power D observed above. The value of D varies depending on the subject (see Table 1). However, all but one of them were found within the range $1 < D < 2$. (We do not agree with Foucault on the value of D.) The inequality $1 < D$ corresponds to the fact that the learning pace should slow down. As for the meaning of the upper bound $D = 2$, we will give a comment later on.

3. Power law in the neural network model

In this section we address an intriguing question whether the neural network model obeys the power law as observed in the human brain performance. To discuss this property we employ BP model with a single hidden layer which is one of the most popular models of the neural network.

The organization of our network is the encoder type (Rumelhart et al. 1986). It consists of N_i input units, N_h hidden units and N_o output units with $N_o = N_{in} > N_h$. We prepare pairs of N_o dimensional vectors $(e_a, t_a)(a = 1, 2, \cdots, M)$ as an aim of the training. The network is trained to display output t_a when it is shown e_a in the input layer.

The learning of the connection weight w_{ij} when it is shown a-th pattern is as usual (Rumelhart, Hinton and Williams 1986):

$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E_a}{\partial w_{ij}} + \alpha \Delta w_{ij}(t), \tag{3.1}$$

where E_a is an error function defined as $E_a = \frac{1}{2}|x_a^{out} - t_a|^2$. Here, the vector x_a^{out} is an N_o dimensional output vector when the network is shown the a-th input vector e_a

and $| \, |$ denotes the usual norm of N_o dimensional vector. The total error is defined as $E = \sum_{a=1} E_a$. The sigmoid function is used to filter the net input on each hidden and output unit as usual. When this network has learned M input-output relations, we say that the network has memorized M items. Thus the learning time in this case is defined as the number of iterations spent in the convergence. This clarifies what are the correspondents of M and t in this simulation.

Since we should measure the learning time universally for the different sessions having the different values of M, we adopt a criterion of convergence such that $\frac{1}{M}E \leq \epsilon$ with ϵ being a fixed small number.

Carrying out the simulation we immediately face the well known problem of BP network model. The time for the learning to converge is sharply dependent on the initial condition of the network weights which are given randomly at the beginning.

Let us consider a distribution of a set of data of the learning time in many sessions. The standard deviation of the learning times is large especially when the size of the network is large and/or M is large. Given such wide distribution of the data of the learning times, which should we call the real learning time of this network?

In order to overcome this problem let us inquire the precise structure of the wide distribution in question. Displyed in Figures 3 (a),(b) and (c) are histograms of the distribution of the learning times for $M = 1$, $M = 3$ and $M = 5$, respectively. These graphs are made from 3000 learning simulations started with different random initial conditions. The structure of the network is 32-16-32 with the parameters $\alpha = 0.75$, $\eta = 0.8$ and $\epsilon = 0.1$.

One apparent feature of this distribution is an existence of a lower cutoff. Any random initial conditions of the network cannot lead to a shorter learning time than that minimum. Close to that minimum there is a sharply increasing peak. The large standard deviation of this distribution resides in the subsequent long tail. This tail shows no exponential decrease. If we compare these histograms, we find that the height of the peak quickly shrinks as M increases. Because the area under these curves is constant, the standard deviation grows rapidly. This means the low reproducibility of the learning time, which is not seen in the human memory.

In order to fill up this gap between BP and the human brain, we propose the following model of the memory system. Suppose that there are many, say $n = 1000$, back propagation networks having the identical structure N_i-N_h-N_o. We assume that these networks are connected to one central neuron and each network sends a singal to the central neuron when it finishes its learning, see Figure 4. Let us see what function this large system has. The large system consists of many subsystems having the structure N_i-N_h-N_o. Suppose that all these subsystems have started learning in parallel way, triggered by an external input information. What happens after the start of learning is readily seen from Figure 3.

Until the minimal learning time is attained, the central neuron receives no signal from the subnetworks. Just after the minimal time, the central neuron receives a plenty of signals, and the signals will pile up, corresponding to the sharp peak seen in the histogram. If the bias of the central neuron is set to an appropriate value, the central neuron sparks by this piling up. We define the learning time of this composite model as the time when the central neuron sparks.

Let us give a more specific definition. One possibility is to define the learning time to be the time of the peak of the histogram in Figure 3. Or we can also define it to be the time at which 20% of the total subnetworks finish their learning. Both definitions specifies the learning time as the time when a sufficient number of subnetworks have completed their learning.

We have no reason to prefer the latter nor have we any reason to prefer the criterion 20% instead of 30% etc. in the latter definition. As we will see later, however, any of these definitions leads to essentially the same bilog plot of t versus M. Thus we adopt the latter definition with the criterion 20% in this paper.

The advantage of this memory model is obvious. We can always finish learning in almost constant time. If we had a single BP network, we are always in the risk of being trapped in a local minimum or starting with a bad initial condition resulting in a very long learning time. The large scale parallel distributed processing of the present model saves the system out of such instability.

Here let us consider the plausibility of our model briefly from the standpoint of the real human brain. The connections from the subsystems to the central neuron are regarded as synaptic connections. A real neuron in the brain in general receives the synaptic connections from $10^2 \sim 10^4$ other neurons (Peters,Palay and Webster 1976). One is not much surprised if there realizes a network of the real neurons, an essential structure of which is the same as our model. The learning time of our composite model is defined to be the time when an ample number of subnetworks have completed their course. One important point to note is that these subnetworks serve backups to the current job. Therefore the learning time we defined is rephrased as the time needed to construct the memory in the brain with high security. It is again a natural idea that our brain gets a concrete memory by such mechanism. If it is dangerous that there is only one central neuron, we simply have to generalize our model as to contain several copies of the central neuron.

Returning to our simulation, now we give our main result of this paper. Figure 5 represents the bilog plot of t versus M of our memory model with the structure 32-5-32 and with the parameters $\alpha = 0.8$, $\eta = 0.75$ and $\epsilon = 0.01$ and the number of subnetworks $n = 100$. The plots fit a line with a good precision. The value of D is also close to that of the human case. We have performed the simulation for many different values of parameters and structure of the network and found that the power law universally holds.

As we mentioned above we define the learning time as the time when 20% of the total subnetworks complete their learning. In Figure 6 we see how the graph change if we take the ratio 30% instead of 20%. This tells us that the choice of the ratio is inessential.

In most cases the value of D seems to be consistent with $D = 2$. What is the meaning of $D = 2$? This value is considered to be the worst bound of D in the following sense.

The slowing down of the pace of the memory is a result of the interferences among different items to memorize. In order for the network to embed M items into itself without mixing, $M(M-1)/2$ tasks will be needed since the number of possible interferences among M items is equal to $M(M-1)/2$ (which is the number of bonds among M points). Accordingly we can roughly estimate the time necessary to complete such job as $t \propto M(M-1)/2 \sim M^2$. This implies $D = 2$. In that sense, Table 1 indicates that the human brain learns items more smartly than such a worst way.

Now we discuss the universality of the value of D in this model. It is verified by the simulation that the change of α and η only lead to the change of $c = t(1)$ and the value of D is unchanged. This is in fact analytically understood.

The learning time of the network scales as

$$t \propto (1-\alpha)/\eta$$

irrespective to the value of M, which we will show in the Appendix. Of course this does not mean that you can make the learning time shorter as you like by taking α close to unity, and/or taking η large. The scaling itself breaks in such an extreme region. Because this scaling applies uniformly in M, the change is only felt by the constant c and D is unchanged at least in the scaling region.

Next we ask how about the criterion parameter ϵ. The change of the value of ϵ of course modifies the learning time. However it is shown by our simulation that this modification is again uniform in M. The universality of D is not violated again. To illustrate it we show a plot with stronger criterion $\epsilon = 0.01$ in Figure 5 and a plot with weaker one $\epsilon = 0.1$ in Figure 6. Moreover the learning time is insensitive to ϵ. In fact while we vary ϵ from 0.005 to 0.5, t changes only by factor 10. From this and the fact that the change of the learning time caused by the change of ϵ is uniform, we know that we do not need

to discuss what value we should choose for ϵ seriously.

The only remaining factor to be considered would be the structure of the network connection. We have at present no theoretical idea how the power law is affected by the change of the structure N_i-N_h-N_o. Most of the result of our simulation seem to be consistent with $D = 2$.

Although the global behavior is naturally understood as $t \propto M^2$, the precise look at the bilog plot shows that there is some vibrating mode around the power law M^2. We do not have any idea on this correction to M^2, but this behavior seems to be universal and may reflect some dynamical structure of the model.

Next we consider the finite size effect of this model. It is known (Rumelhart et al. 1986)that the upper limit of the memory M of the encoder network is estimated as $\sim 2^{N_H}$ since the binary encoding is used on the hidden layer. Therefore the power law should break in the vicinity of the upper limit of M. We examined this finite size effect using a small network 10-3-10 whose limit of M is eight. Figure 7 shows the corresponding deviation from the power law.

Now we focus our attention on the value of c which is the learning time to memorize one item. By identifying the power law of the human brain and that of BP model, we can determine the relation between the time t (second) of the human experiment and the time t (iteration) of BP. In a large parameter region we find that a unit learning corresponds to $o(1) \sim o(10)$ msec. It is interesting that this reaction time is physiologically reasonable (Feldman 1982,1985) (Lynch,Mountcastle,Talbot and Yin 1977).

4. Discussion

In this paper we have reexamined the power law of the human brain proposed by Foucault and found it positive. Furthermore we have found out that BP model also obeys the same power law and that the values of exponent in cases are close to each other. By identifying the power laws of both we roughly estimated a correspondence between the human time and the time of BP. This leads to a physiologically reasonable value of the time spent in one step of learning for large ranges of the parameters. It is of great importance that even very simplified model like BP exhibits quantitatively the same performance with the real human brain. This result will suggest that the study of the neural network model has an ample possibility to cast the light on the secret of the human brain.

We pointed out that the stability of the learning time is a gap between the human brain and the single BP network and proposed the composite model which fills up this gap in technically and physiologically natural way.

Considering the value of D, it seems that D's of the human brain are smaller than that of BP. On this problem we are going to study what kind of modification of the model leads to the value of D closer to that of the human brain. This approach is expected to clarify the difference between the human brain and BP. For example we can show that if we change the stopping condition of the learning from the criterion $E/M < \epsilon$ to $E_a < \epsilon$ for all a, the exponent comes in the range $2 < D < 3$. Unfortunately this modification makes the model further from the human brain.

It is also a charming problem to derive the power law analytically. As to this problem we are trying on the following line. We define N_o-dimensional vectors of error as $e_a = t_a - x_a^{out}$ for $a = 1, 2, \cdots, M$. The time development of these are determined by that of w_{ij}, and the time development of w_{ij} is controlled by the learning equation (3.1). Especially when $\alpha = 0$ we obtain the following differential equation for e,

$$\partial_t e = -\eta B e$$

where B is a symmetric matrix of the size $MN_o \times MN_o$ which depends on w_{ij}. This appearance of η is nothing but the reason why the learning time scales as $\propto 1/\eta$. Similarly, in order to study the M dependence of the learning time, we should examine what kind of scaling can appear when the size of the matrix B changes as M. This approach is now under investigation.

APPENDIX

In this Appendix we show that the learning time of the BP network scales as $(1-\alpha)/\eta$.

We consider the error of the network output, which is a function of time t with the parameters α and η. If we denote this function as $e(t,\eta,\alpha)$, the learning time t_0 is determined as $e(t_0,\eta,\alpha)=\epsilon$. When the momentum coefficient α vanishes, the learning equation (3.1) is written in a continuum limit as

$$\frac{dw_{ij}}{dt} = -\eta \frac{\partial E_a}{\partial w_{ij}}.$$

This means that the learning curve $e(t,\eta,0)$ is a function of ηt, namely we get $e(t,\eta,0)=e(\eta t)$. On the other hand when $\alpha \neq 0$ we can rewrite (3.1) as

$$\frac{d^2 w_{ij}}{dt^2} = -\kappa \frac{dw_{ij}}{dt} - \eta \frac{\partial E_a}{\partial w_{ij}},$$

with $\kappa = 1-\alpha$. This is nothing but the Newton's equation of motion with the potential ηE_a and with a friction coefficient κ. If we divide the equation by κ^2, we obtain to the same form of the equation with the replacements, $\kappa \to 1$,$\eta \to \eta/\kappa^2$ and $t \to \kappa t$. Since $\kappa = 1$ means $\alpha = 0$, this teaches us that $e(t,\eta,\alpha) = e(\kappa t, \eta/\kappa^2, 0)$.

Combining the above $e(t,\eta,0)=e(\eta t)$ we finally get the following scaling relation:

$$e(t,\eta,\alpha) = e\left(\frac{\eta}{\kappa} t\right).$$

which signifies that the learning time scales as $t \propto \kappa/\eta = (1-\alpha)/\eta$. This scaling property is certainly verified by our simulation.

NOMENCLATURE

M: The number of items to memorize.

$t(M)$: The time necessary to memorize M items.

D: The exponent of the power law of the memory performance.

c: The constant of the proportionality appearing in the power law.

n: The number of subnetworks of our memory model.

N_i, N_h, N_o: The numbers of input, hidden and output units, respectively.

t_a: The a-th target signal.

x_a^{out}: The a-th output.

w_{ij}: The network connection of the back propagation network.

t: The continuous variable parametrizing the learning step of the back propagation network model.

α, η: The momentum coefficient and learning coefficient of the learning algorithm of the back propagation network model.

E_a, E: The error at the output layer when the network is shown a-th input, the total error.

ϵ: The criterion of the error to stop the learning.

κ: is equal to $1 - \alpha$.

e_a: The a-th error of the model ($= t_a - x_a^{out}$).

B: The symmetric matrix driving the time development of the error of the back propagation network with the vanishing momentum coefficient.

REFERENCES

Amari,S. (1967), IEEE Trans. EC, **16**,no.3,299

Câteau,H., Fuchikami,N., Nakajima,T. and Nunokawa,H.(1992), IEICE Technical report **91** 87 (in Japanese)

Feldman,J.A. (1982), Cognitive Science **6**, 205

Feldman,J.A. (1985), Cognitive Science **9**, 1

Foucault,M. (1913), Anée,Psychol.**19**, 218

Lynch,J.C.,Mountcastle,V.B.,Talbot,W.H. and Yin,T.C.T.(1977), Journal of Neurophysiology,**40** 362

McCulloch,W.S. and Pitts,W.H. (1943), Bull.Math.Biophys.**5**,115

Nakajima,T., Fuchikami,N., Câteau,H. and Nunokawa,H.(1992) , Power law in the performance of the human memory and a simulation with a neural network model, to appear in proceedings of ISKIT'92

Peters,A.,Palay,S.L. and Webster,H.deF. (1976) *The fine structure of the nervous system.*,Philadelphia:W.B.Saunders.

Rumelhart,D.E., Hinton,G.E. and Williams,R.J. (1986), Nature **323**,533

Rumelhart,D.E., McClelland,J.L. and the PDP research Group (1986), *Parallel Distributed Processing vol.I & II*,MIT Press, Cambridge

Widrow,B. and Hoff,M.E. (1960), Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record Part **4**,94

Table Legends

Table 1. These two tables (a) and (b) represent the result of the psychological experiment. The first column shows the names of the subjects, the second column shows how many times the subject is checked his/her memory. The third and fourth columns show the constants c and D appearing in the power law (eq. (2.1) in the text). The last two columns show the significance levels of the linearity of the plots under the assumption that the standard deviation of the single plot point is 10% and 20% of its value, respectively.

Figure Legends

Figure 1. The bilog plot of the time versus the number of memory. Data from ordinary people.

Figure 2. The bilog plot of the time versus the number of memory. Data from a talent.

Figure 3. The histogram of the learning times made by 3000 learning sessions. The structure of the network is 32-16-32, the parameters are taken as $\alpha = 0.75$, $\eta = 0.8$ and $\epsilon = 0.1$. Figures (a), (b) and (c) represent the cases $M = 1$, $M = 3$ and $M = 5$ respectively.

Figure 4. The picture representing our memory model. Each small circle arranged downward represents the back propagation network of the encoder type. Each one sends signal upward to the central neuron when it finish its learning.

Figure 5. The plots with common logarithm of t versus M of the back propagation model which is implemented as our memory model. The structure of the network is 32-5-32, the number of subnetworks is 100, the parameters are taken as $\alpha = 0.8$, $\eta = 0.75$ and $\epsilon = 0.01$. c and D are defined as eq. (2.1) in the text.

Figure 6. The bilog plots with different criterion value which corresponds to the bias of the central neuron. In this case $\epsilon = 0.1$ and other network conditions are the same as in Figure 5. (a) and (b) represent the 20% and 30% case respectively.

Figure 7. The bilog plot which represents the deviation from the power law. The structure of the network is 10-3-10 and the number of subnetworks is 100. The parameters are the following: $\alpha = 0.75$, $\eta = 0.8$ and $\epsilon = 0.1$. We can see the deviation from the line at $M = 9$.

Table 1.

(a) People with abnormally high memory

name	No. of data	c	D	S. L. (%) (10% error)	S. L. (%) (20% error)
1. Dia Mandi	8	0.68	1.8	0.2	53
2. Ishihara	8	0.97	1.2	8.8	85

(b) People with ordinary memory

name	No. of data	c	D	S. L. (%) (10% error)	S. L. (%) (20% error)
1. T. S.	5	0.70	1.9	91	99
2. S. K.	4	0.92	1.5	95	99
3. I. K.	5	0.96	1.3	83	97
4. A. N.	5	0.75	1.6	39	86
5. T. K	5	0.44	1.9	35	84
6. M. M.	5	1.0	1.7	65	94
7. K. I.	4	2.3	1.2	56	87
8. T. N.	5	1.1	1.5	1.4	45
9. K. K.	4	5.0	1.1	6.4	51
10. Y. K.	5	1.3	1.3	54	91
11. S. H.	4	0.16	3.9	86	98
12. S. S.	4	1.3	1.5	11	69
13. K. Y.	4	0.73	1.5	86	96
14. H. N.	4	0.60	1.9	90	98
average		1.23	1.70	57.1	85.3