Mj. Int. J. Sci. Tech. **2008**, *2*(01), 210-226

Maejo International
Journal of Science and Technology
ISSN 1905-7873
Available online at www.mijst.mju.ac.th

Full Paper

Quantitative characterisation of an engineering write-up using random walk analysis

Babatunde Alabi[1], Tajudeen A. Salau[2], and Sunday A. Oke[3,*]

[1] Federal University of Petroleum Resources, Effurun, Delta State, Nigeria

[2] Department of Mechanical Engineering, University of Ibadan, Nigeria

[3] Department of Mechanical Engineering, University of Lagos, Nigeria

* Author to whom correspondence should be addressed, e-mail: sa_oke@yahoo.com

Received: 6 November 2007 / Accepted: 21 February 2008 / Published: 28 February 2008

Abstract: This contribution reports on the investigation of correlation properties in an English scientific text (engineering write-up) by means of a random walk. Though the idea to use a random walk to characterise correlations is not new (it was used e.g. in the genome analysis and in the analysis of texts), a random walk approach to the analysis of an English scientific text is still far from being exploited in its full strength as demonstrated in this paper. A method of high-dimensional embedding is proposed. Case examples were drawn arbitrarily from four engineering write-ups (Ph.D. synopsis) of three engineering departments in the Faculty of Technology, University of Ibadan, Nigeria. Thirteen additional analyses of non-engineering English texts were made and the results compared to the engineering English texts. Thus, a total of seventeen write-ups of eight Faculties and sixteen Departments of the University of Ibadan were considered. The characterising exponents which relate the average distance of random walkers away from a known starting position to the elapsed time steps were estimated for the seventeen cases according to the power law and in three different dimensional spaces. The average characteristic exponent obtained for the seventeen cases and over three different dimensional spaces studied was 1.42 to 2-decimal with a minimum and a maximum coefficient of determination (R^2) of 0.9495 and 0.9994 respectively. This is found to be 284% of the average characterising exponent value (0.5), as supported by the literature for random walkers based on the pseudo-random number generator. The average characteristic exponent obtained for the four cases that were engineering-based and over the three different dimensional studied spaces was 1.41 to 2-decimal (closer by 99.3% to 1.42) with a minimum and a maximum coefficient of determination (R^2) of 0.9507 and 0.9974 respectively. This is found to be 282% of the average characterising exponent value (0.5), as supported by the literature for random walkers based on the pseudo-random number generator. In view of the range of the average characterising exponent across Faculties and the closeness of the average characterising exponent in the engineering-based cases in particular, it can be concluded that the engineering writing is strongly correlated. This study recommends that a very high characterising exponent value (e.g 1.42) is a mark of a very good engineering write-up.

Keywords: engineering write-up, random walker analysis, quantitative evaluation, power law

---

Introduction

In the evaluation of scholarly publications, a major worldwide challenge has been the need to eliminate subjectivity in evaluating papers in terms of contribution to knowledge and the overall suitability for publication [1]. Manuscripts for publication are supposed to be objectively evaluated for presentation, scientific merit, quality of English, spellings and grammar, and the overall quality [2]. These measurement parameters also apply to the evaluation of projects, dissertations, and Ph.D. theses. Unfortunately, it is a common experience that some evaluations are biased, with judgements made unfairly in many instances. Evaluators are often accused of race discrimination, hatred, and inferiority complex. Their evaluation decisions and comments may have been motivated by the qualitative means of evaluation for most scientific writings, which depend on personal judgements.

Faced with this challenge, the scientific community has been making efforts at solving this important problem. An attractive approach is the development of a quantitative methodology with which all works may be evaluated. It is expected that the use of such a methodology should produce similar results from the assessments of two independent reviewers. Thus, engineering write-ups, which may include papers and Ph.D. theses, should be quantitatively evaluated. The aim of this study is to develop and apply a quantitative approach in evaluating engineering write-ups. Specifically, this contribution investigates the correlation properties in an English scientific text by means of a random walk. The framework of this approach is built on the principles supporting random walker movements in three different dimensional spaces (1-D, 2-D and 3-D) [3-5].

The idea of using a random walk to characterise correlations is not new. It was used e.g. in the genome analysis by Peng et al. [6] and Buldyrev et al. [7]. The former group analysed the fluctuations in the correlation exponents obtained for noncoding DNA sequences by generating correlated random sequences of comparable length and studied the fluctuations in this control system. They concluded that the DNA exponent fluctuations are consistent with those obtained from the control sequences having long-range power-law correlations. The latter researchers proposed a generalised Lévy walk (GLW) to model fractal landscapes observed in noncoding DNA sequences and provided an approximation of genomic DNA sequences, such as the distribution of strand-biased regions (those with an excess of one type of nucleotide) as well as local changes in the slope of the correlation exponent α. Random walk concept has also been applied in the analysis of texts [8-9]. For example, Schinner [9] analysed the Voynich manuscript using random walk mapping and token/syllable repetition statistics. The results significantly tighten the boundaries for possible interpretations by suggesting that the text has been generated by a stochastic process rather than by encoding or encryption of the language.

A major step in the approach of random walk is to analyse the empirical relationships between the average distance of random walkers from a known starting position (datum) and time steps [10-11]. In addition, a comparative analysis of the empirical relationship obtained with values cited in the literature for random walkers is made in the present work. This analysis is driven by the pseudo-random number generator in three different dimensional spaces. It is hoped that the findings of this study will stimulate further research and enable the research community to know whether or not an arbitrary engineering write-up is a form of random process, and that readers and researchers might be enlightened on how to have a measure of the relative quality of the write-up quantitatively.

Methodology

The methodology presented here is hinged on the random walker movement (RWM) principles, which are driven by the pseudo-random number generator in any space (1-D, 2-D or 3-D). RWM is governed empirically as [12]:

$$D_{av} \propto T^{\beta} \qquad (1)$$

Expression (1) states that D_{av} varies directly as the β^{th} power of T. Here, β is referred to as the characteristic exponent. D_{av} is termed as the average distance of a number of random walk from a known starting position (datum after a known time step has elapsed), and T is the time step. It should be noted that for a particle undergoing pure random motion with reference to a fixed starting position in 1-D, 2-D or 3-D Euclidean space, the characterising exponent β is equal to 0.5. The explanation is that the particle moves away from the fixed starting position such that its average distance (D_{av}) is directly proportional to the square root of elapsed time (T). Now let us imagine the same particle being driven similarly ("randomly" so to say) by the organised string of the English language alphabets of equally organised thoughts called the "Engineering Write-ups", the average distance (D_{av}) measured is expected to be higher for the same elapsed time.

The intent of the author of the "Engineering Write-ups" is foremost to communicate the whole of his conception to the readers in the shortest possible time. This situation may be likened to a very fast moving particle aimed at a target. To establish the corresponding expression (1) for random walk movement being driven by an engineering write-up in any space, the method description is as follows:

> Let us assume a 10-letter word "Technology" in an engineering write-up which involves three random walks. In a 1-dimensional space there are only two possible random step movements (forward or backward). These are labelled 1 or 2 for digital identification purpose. In a 2-dimensional space there are only four possible random step

movements: two in each of the two mutually perpendicular directions. These are labelled 1, 2, 3 or 4 for digital identification purpose. In a 3-dimensional space there are only six random step movements: two in each of the three mutually perpendicular directions. These are labelled 1, 2, 3, 4, 5 or 6 for digital identification purpose.

Based on the description made above, a serial identification of the English alphabets without differentiation between small and capital letters is made. Each of the alphabets is then matched to a serial number. For example, "A" or "a" is matched to "1", "B" or "b" to "2", "C" or "c" to "3", … , and "Z" or "z" to "26" (Table 1). Thus, the interpretation of the word "Technology" is shown in Table 2.

Table 1. Serial identification of English alphabets without differentiation between small and capital letters

Alphabet	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Serial Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

In Table 2, the interpretation of Mod(SN,2) is that "SN" is divided by two (2) and the modulus of the remainder is obtained. This definition is similar for the 2-D and 3-D situations. It is assumed that the three random walks used have one consecutive time step delay interval before starting and ending respectively. Thus random walk 1 starts to be active on the instruction of the first letter (T) of the word "Technology" and stops to be active on the third letter (C) for an experiment involving three random steps. Random walk 2 becomes active on the instruction of the second letter (E) and stops to be active on the fourth letter (H) for an experiment involving three random steps. Random walk 3 becomes active on the instruction of the third letter (C) and stops to be active on the sixth letter (N) for an experiment involving three random steps. Referring to Table 2, the random direction followed by the walker depends on the dimensional space being investigated. The interpretation of Table 2 may be made from an understanding of expression (1). Simply explained, it means that the distance of each walker from the known starting position can be measured. Measurement is done at the end of the specified time steps and at the end of a constant step size. In the evaluation of the write-ups, it may be necessary to statistically determine the relationship among the sets of solutions. This is pursued in this work with the aid of the correlation coefficient (R) between two sets of data, and the coefficient of determination (R^2), as stated in the section on results.

Table 2. Digital random movement in three different dimensional spaces of the word "Technology"

Alphabets, Space /Instructions	10-Letters One Word "Technology"									
	T	E	C	H	N	O	L	O	G	Y
Alphabet Serial Number (SN)	20	5	3	8	14	15	12	15	7	25
1-D: Mod(SN,2)+1	1	2	2	1	1	2	1	2	2	2
2-D: Mod(SN,4)+1	1	2	4	1	3	4	1	4	4	2
3-D: Mod(SN,6)+1	3	6	4	3	3	4	1	4	2	2

Case Study

Four selected cases of engineering write-ups (synopses of abstracts) were made from the Postgraduate School, University of Ibadan, Ibadan, Nigeria. The synopses of abstracts (Ph.D. Theses 2000-2002) involved three Engineering Departments (Agriculture, Food Technology and Mechanical). These served as entry data to a FORTRAN-based algorithm for analysing the data for the purpose of establishing the corresponding expression (1) in each of the four cases studied. That is, each of these data was analysed like it was previously described for the 10-letter word "Technology". However, one hundred random walks (100), five hundred and fifty time steps (550) and constant step size (unity) were used to study each of the four cases. The numbers of alphabets used to compose the four write-ups were 628, 820, 884, and 682. Additional thirteen cases (13) involving different Departments were drawn arbitrarily for the same analysis above, using the same FORTRAN platform across another seven (7) Faculties. The findings of this study are presented in tabular as well as graphical forms in the results section.

Results

After extensive laboratory experiments, eight additional tables were generated. The first relates to the sample data for 100 random walks in 1-dimensional space (Table 3). In Table 4, the results of the sample logarithm of data for 100 random walks in 1-dimensional space are presented. Table 5 shows the sample data for 100 random walks in 2-dimensional space. An additional table, Table 6, was generated which contains the sample logarithm of data for 100 random walks in 2-dimensional space. The results presented in Table 7 relate to the sample data for 100 random walks in 3-dimensional space. Table 8 shows the sample logarithm of data for 100 random walks in 3-dimensional space. In Table 9, the characterising exponents of random walk movement in three different dimensional spaces are displayed. This employs four different engineering write-ups.

Table 10 shows R^2 of line of best fit for the four engineering write-up cases and three dimensional spaces. The graphs show the log of average distance against the log of time steps for the four cases in the three different dimensional spaces (Figures 1 to 12). In addition, Figure 13 shows the results of the first random walk (constant step size) trace for case 4 in 2-D. From Table 11, the Faculty of Public Health is the youngest of all studied Faculties of the University and this may explain why only one Department exists in the documentation by Olayinka et al. [4]. In Table 12, the average characterising exponent for cases in the Faculty of Technology is 1.4155 while that for all Faculties studied is 1.42 to 2-decimal.

Table 3. Sample data for 100 random walks in 1-dimensional space

Time step	Average distance from starting point (D_{av})			
	Case 1	Case 2	Case 3	Case 4
1	1.00	1.00	1.00	1.00
2	1.32	1.28	1.34	1.36
3	2.32	2.28	2.34	2.36
4	2.86	2.72	2.96	2.96
5	3.90	3.74	4.04	4.04
6	4.66	4.32	4.90	4.76
7	5.74	5.40	6.10	5.88
8	6.60	6.12	7.14	6.76
9	7.74	7.22	8.48	7.90
10	8.72	8.02	9.66	8.82
11	9.92	9.18	11.10	9.98
12	10.94	10.10	12.38	10.76
13	12.16	11.32	13.94	11.94
14	13.14	12.28	15.30	12.74
15	14.34	13.56	16.92	13.94
16	15.28	14.60	18.36	14.82
17	16.44	15.94	20.08	16.02
18	17.30	16.98	21.64	16.90
19	18.46	18.36	23.42	18.14
20	19.28	19.42	25.12	19.02

Table 4. Sample logarithm of data for 100 random walks in 1-dimensional space

Log of time step	Log of average distance from starting point {Log (D_{av})}			
	Case 1	Case 2	Case 3	Case 4
0.00000	0.00000	0.00000	0.00000	0.00000
0.69315	0.27763	0.24686	0.29267	0.30748
1.09861	0.84157	0.82418	0.85015	0.85866
1.38629	1.05082	1.00063	1.08519	1.08519
1.60944	1.36098	1.31909	1.39624	1.39624
1.79176	1.53902	1.46326	1.58924	1.56025
1.94591	1.74746	1.68640	1.80829	1.77156
2.07944	1.88707	1.81156	1.96571	1.91102
2.19722	2.04640	1.97685	2.13771	2.06686
2.30259	2.16562	2.08194	2.26799	2.17702
2.39790	2.29455	2.21703	2.40695	2.30058
2.48491	2.39243	2.31254	2.51608	2.37584
2.56495	2.49815	2.42657	2.63476	2.47989
2.63906	2.57566	2.50797	2.72785	2.54475
2.70805	2.66305	2.60712	2.82850	2.63476
2.77259	2.72654	2.68102	2.91017	2.69598
2.83321	2.79972	2.76883	2.99972	2.77384
2.89037	2.85071	2.83204	3.07454	2.82731
2.94444	2.91561	2.91017	3.15359	2.89812
2.99573	2.95907	2.96630	3.22366	2.94549

Table 5. Sample data for 100 random walks in 2-dimensional space

Time step	Average distance from starting point (D_{av})			
	Case 1	Case 2	Case 3	Case 4
1	1.00000	1.00000	1.0000	1.00000
2	1.86610	1.79196	1.82610	1.83439
3	2.97735	2.79196	2.83846	2.83439
4	3.84051	3.38794	3.61101	3.60693
5	5.12614	4.52391	4.84586	4.85415
6	6.33640	5.48172	6.02711	5.97540
7	7.67573	6.61297	7.30488	7.22845
8	9.06319	7.64927	8.63354	8.54204
9	10.58608	8.92885	10.17380	10.05175
10	12.01359	10.08524	11.69568	11.52777
11	13.55648	11.42482	13.32087	13.09810
12	15.12586	12.77247	15.01240	14.69195
13	16.80845	14.23856	16.77805	16.43040
14	18.38995	15.67711	18.54148	18.14368
15	20.12937	17.22049	20.42777	20.01306
16	21.84876	18.74635	22.32399	21.89820
17	23.65178	20.39897	24.29166	23.85775
18	25.39287	21.98258	26.28230	25.85255
19	27.28847	23.69270	28.36809	27.96545
20	29.15477	25.32019	30.44992	30.09641

Table 6. Sample logarithm of data for 100 random walks in 2-dimensional space

Log of time step	Log of average distance from starting point {Log (D_{av})}			
	Case 1	Case 2	Case 3	Case 4
0.00000	0.00000	0.00000	0.00000	0.00000
0.69315	0.62385	0.58331	0.60218	0.60671
1.09861	1.09103	1.02674	1.04326	1.04183
1.38629	1.34561	1.22022	1.28399	1.28286
1.60944	1.63435	1.50938	1.57813	1.57983
1.79176	1.84631	1.70142	1.79627	1.78765
1.94591	2.03806	1.88903	1.98854	1.97802
2.07944	2.20422	2.03461	2.15565	2.14500
2.19722	2.35954	2.18929	2.31982	2.30775
2.30259	2.48604	2.31107	2.45922	2.44476
2.39790	2.60686	2.43579	2.58933	2.57247
2.48491	2.71641	2.54729	2.70888	2.68730
2.56495	2.82188	2.65595	2.82007	2.79913
2.63906	2.91180	2.75220	2.92001	2.89832
2.70805	3.00218	2.84610	3.01690	2.99639
2.77259	3.08414	2.93100	3.10566	3.08640
2.83321	3.16344	3.01548	3.19013	3.17211
2.89037	3.23447	3.09025	3.26890	3.25241
2.94444	3.30646	3.16517	3.34526	3.33097
2.99573	3.37262	3.23160	3.41608	3.40441

Table 7. Sample data for 100 random walks in 3-dimensional space

Time step	Average distance from starting point (D_{av})			
	Case 1	Case 2	Case 3	Case 4
1	1.00000	1.00000	1.00000	1.00000
2	1.88267	1.83439	1.80610	1.86267
3	3.04144	2.97852	2.91591	3.02372
4	4.16210	4.04124	3.97414	4.13716
5	5.37859	5.21609	5.07255	5.30697
6	6.52545	6.31775	6.15279	6.41242
7	7.90703	7.74161	7.45029	7.76700
8	9.32355	9.16919	8.66822	9.06876
9	10.85630	10.74879	10.05908	10.45262
10	12.41551	12.34331	11.38322	11.84222
11	13.97233	14.04416	12.81896	13.27477
12	15.59779	15.86415	14.30008	14.71744
13	17.25155	17.80560	15.84431	16.24775
14	18.92464	19.82249	17.36775	17.74230
15	20.62928	21.90694	18.95914	19.30189
16	22.31717	24.07942	20.50274	20.76793
17	24.04298	26.38102	22.16668	22.38742
18	25.75615	28.76229	23.80007	23.92553
19	27.51095	31.25052	25.53355	25.56330
20	29.23524	33.78845	27.23595	27.18490

Table 8. Sample logarithm of data for 100 random walks in 3-dimensional space

Log of time step	Log of average distance from starting point {Log (D_{av})}			
	Case 1	Case 2	Case 3	Case 4
0.00000	0.00000	0.00000	0.00000	0.00000
0.69315	0.63269	0.60671	0.59117	0.62201
1.09861	1.11233	1.09143	1.07018	1.10649
1.38629	1.42602	1.39655	1.37981	1.42001
1.60944	1.68243	1.65175	1.62384	1.66902
1.79176	1.87571	1.84336	1.81691	1.85824
1.94591	2.06775	2.04661	2.00825	2.04988
2.07944	2.23254	2.21585	2.15966	2.20484
2.19722	2.38475	2.37479	2.30848	2.34685
2.30259	2.51895	2.51311	2.43214	2.47167
2.39790	2.63708	2.64221	2.55093	2.58587
2.48491	2.74713	2.76406	2.66027	2.68903
2.56495	2.84790	2.87951	2.76281	2.78795
2.63906	2.94046	2.98682	2.85461	2.87595
2.70805	3.02671	3.08680	2.94229	2.96020
2.77259	3.10536	3.18136	3.02056	3.03341
2.83321	3.17984	3.27264	3.09859	3.10850
2.89037	3.24867	3.35907	3.16969	3.17495
2.94444	3.31458	3.44204	3.23999	3.24116
2.99573	3.37537	3.52012	3.30454	3.30266

Table 9. Characterising exponents of random walk movement in three different dimensional spaces using 4 different engineering write-ups

Number of dimensional space	Characterising exponent			
	Case 1	Case 2	Case 3	Case 4
1	1.7157	1.2165	1.3139	1.1130
2	1.4577	1.2462	1.4753	1.2748
3	1.5040	1.6084	1.4723	1.5885

Table 10. R^2 of line of best fit for 4 engineering write-up cases in three different dimensional spaces

Number of Dimensional Space	R^2			
	Case 1	Case 2	Case 3	Case 4
1	0.9507	0.9962	0.9877	0.9974
2	0.9832	0.9887	0.9950	0.9899
3	0.9767	0.9959	0.9904	0.9866

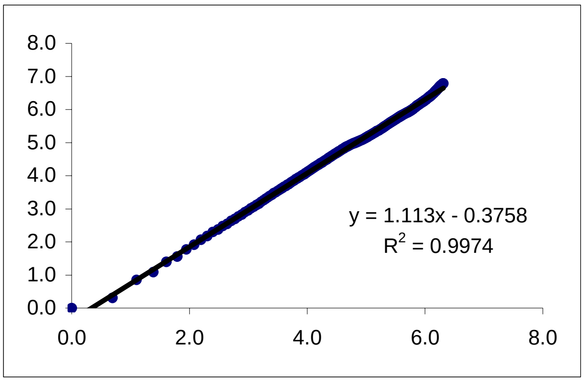

Figure 1. Log of average distance versus log of time step for case 1 in 1-D

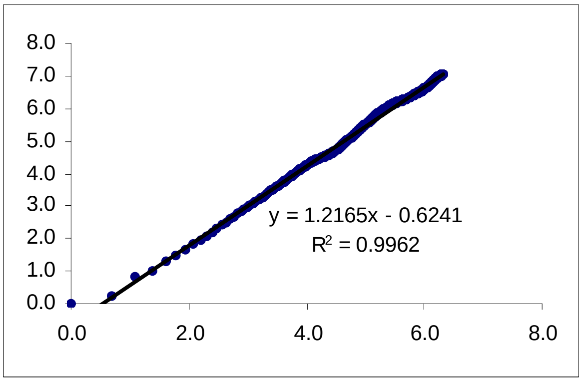

Figure 2. Log of average distance versus log of time step for case 2 in 1-D

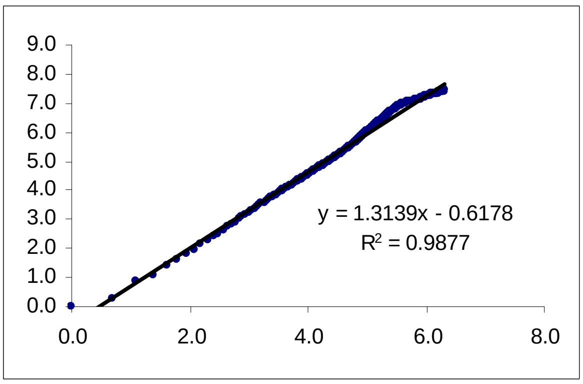

Figure 3. Log of average distance versus log of time step for case 3 in 1-D

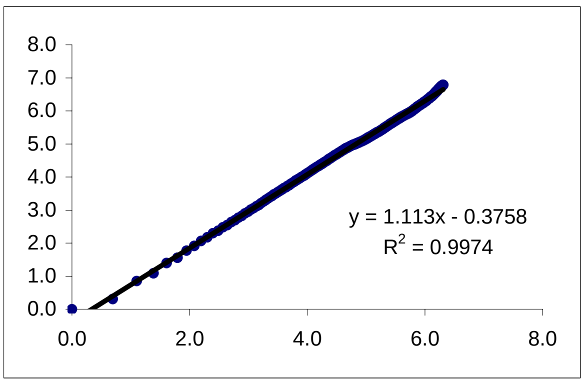

Figure 4. Log of average distance versus log of time step for case 4 in 1-D

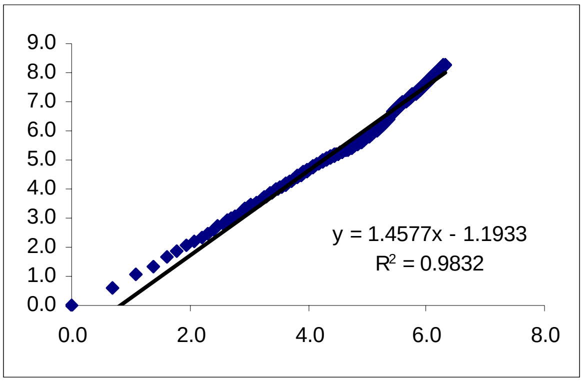

Figure 5. Log of average distance versus log of time step for case 1 in 2-D

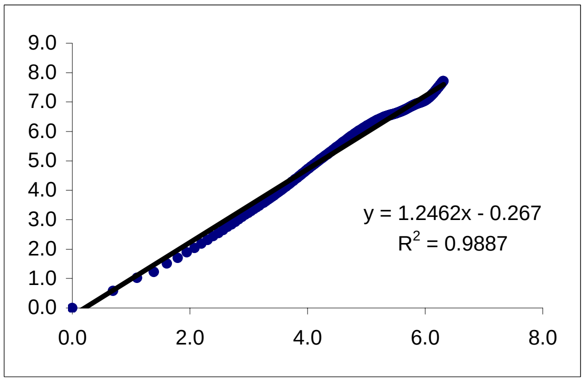

Figure 6. Log of average distance versus log of time step for case 2 in 2-D

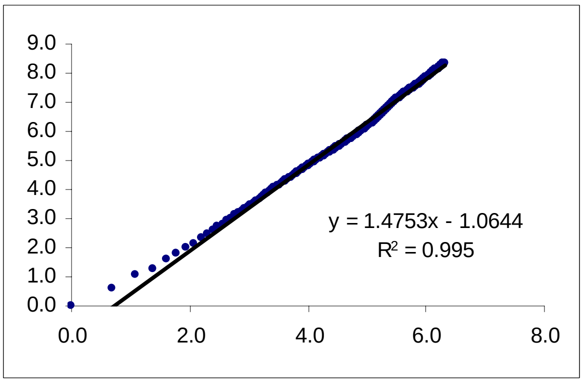

Figure 7. Log of average distance versus log of time step for case 3 in 2-D

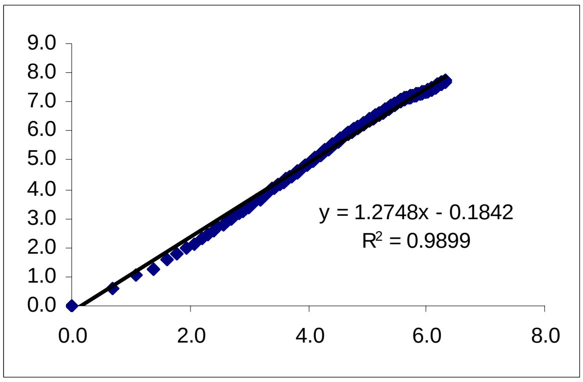

Figure 8. Log of average distance versus log of time step for case 4 in 2-D

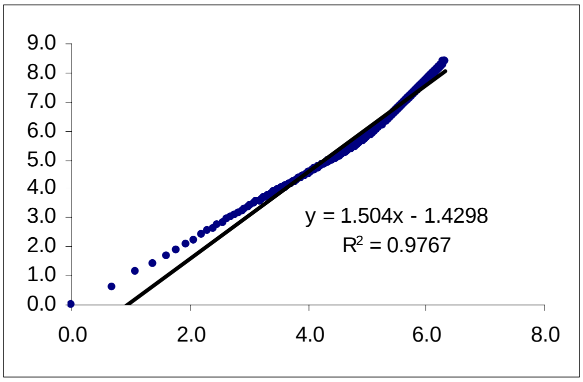

Figure 9. Log of average distance versus log of time step for case 1 in 3-D

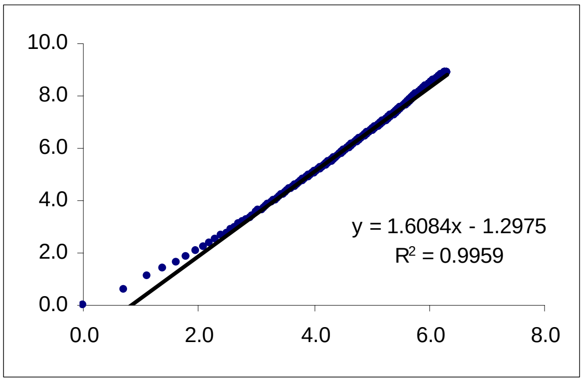

Figure 10. Log of average distance versus log of time step for case 2 in 3-D

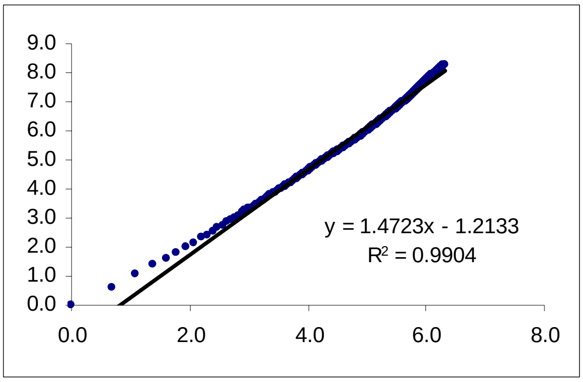

Figure 11. Log of average distance versus log of time step for case 3 in 3-D

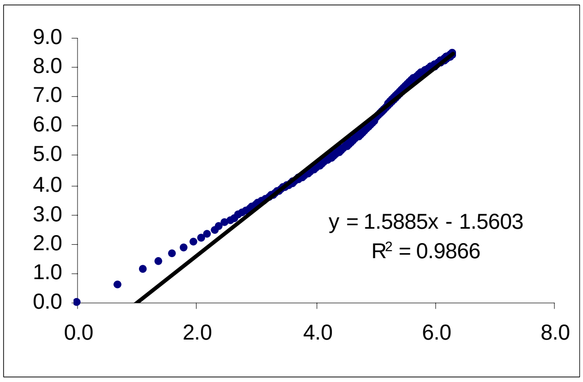

Figure 12. Log of average distance versus log of time step for case 4 in 3-D

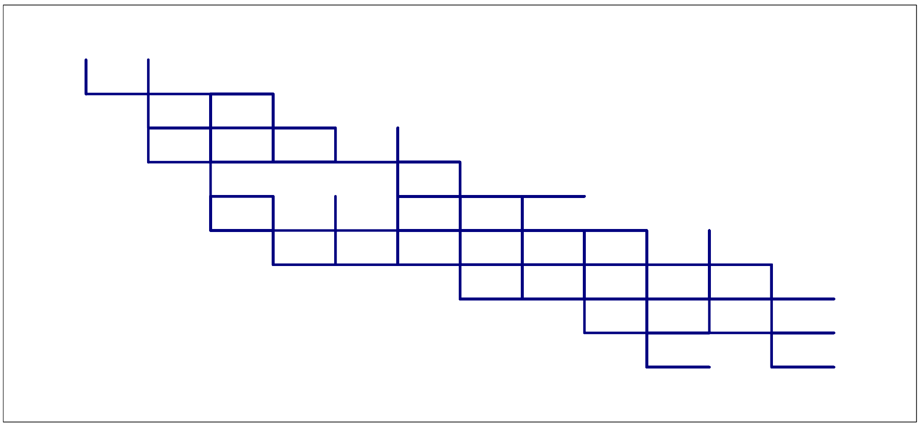

Figure 13. First random walk (constant step size) trace for case 4 in 2-D

Table 11. Characterising exponents of random walk movement and R^2 of line of fit for additional thirteen cases of write-ups

Faculty	Department	Number of alphabets used	1-Dimension space		2-Dimension space		3-Dimension space		Average slope (Faculty)	R^2 range (Faculty)
			Expo-nent	R^2	Expo-nent	R^2	Expo-nent	R^2		
Agriculture & Forestry	Agricultural Economics	891	1.3482	0.9949	1.7260	0.9794	1.4600	0.9962	1.4129	0.9794 to 0.9962
	Wildlife & Fisheries Management	834	1.2455	0.9941	1.2984	0.9931	1.3994	0.9838		
Basic Medical Sciences	Biochemistry	618	1.3980	0.9495	1.3323	0.9849	1.6797	0.9732	1.3796	0.9495 to 0.9953
	Virology	612	1.1560	0.9799	1.2606	0.9924	1.4512	0.9953		
Clinical Sciences & Dentistry	Epidemiology, Medical Statistics & Environmental Health	770	1.1724	0.9936	1.1827	0.9978	1.5425	0.9886	1.3447	0.9831 to 0.9981
	Physiotherapy	606	1.2241	0.9831	1.5546	0.9914	1.3920	0.9981		
Public Health	Epidemiology, Medical Statistics & Environmental Health	864	1.6986	0.9929	1.5508	0.9947	1.6361	0.9953	1.6285	0.9929 to 0.9953
Pharmacy	Pharmaceutical Chemistry	757	1.4429	0.9765	1.4457	0.9944	1.5726	0.9962	1.4846	0.9765 to 0.9962
	Pharmaceutical Microbiology & Clinical Pharmacy	723	1.4159	0.9905	1.4110	0.9861	1.6193	0.9958		
Science	Botany & Microbiology	707	1.2365	0.9900	1.2314	0.9942	1.5396	0.9863	1.3532	0.9863 to 0.9968
	Zoology	897	1.3377	0.9931	1.2544	0.9968	1.5196	0.9927		
Veterinary Medicine	Veterinary &Pharmacology	711	1.3833	0.9931	1.2817	0.9873	1.2756	0.9994	1.3424	0.9873 to 0.9994
	Veterinary Medicine	1097	1.2800	0.9962	1.2956	0.9961	1.5379	0.9922		

Table 12. Characterising exponents of random walk movement and R^2 of line of fit for Faculties

Faculty	Average characterising exponents (Faculty)	R^2 Range (Faculty)
Agriculture & Forestry	1.4129	0.9794 to 0.9962
Basic Medical Sciences	1.3796	0.9495 to 0.9953
Clinical Sciences & Dentistry	1.3447	0.9831 to 0.9981
Public Health	1.6285	0.9929 to 0.9953
Pharmacy	1.4846	0.9765 to 0.9962
Technology	1.4155	0.9507 to 0.9974
Science	1.3532	0.9863 to 0.9968
Veterinary Medicine	1.3424	0.9873 to 0.9994
Average characterising exponents across Faculties of study	1.42 to 2-decimal	

Discussion

Referring to the twelve graphs of log of average distance versus log of time step (four graphs to each corresponding dimensional space) and to Table 9 above, the range of the estimated exponent is found to be between 1.1130 and 1.7157 for all cases and numbers of the dimensional space considered while the average exponent is 1.4155 to four decimal places. The exponents estimated for case 3 in the three-dimensional spaces are fairly identical compared to other cases studied. It is very important to note here that the average exponent found in the literature for random walk movement in three different dimensional spaces with constant step size and using the random number generator as a driver is 0.5. Referring to Table 10 above, the range of the estimated R^2 is found to be between 0.9507 and 0.9974 for all cases and numbers of dimensional space considered, with case 1 observed in 1-dimensional space having a minimum R^2 value of 0.9507.

Results in Table 12 involving eight Faculties show that the average characterising exponent range is between 1.3424 and 1.6285. Experimentation with all seventeen cases over three different dimensional spaces shows that the average characteristic exponent is 1.42. The analysis of the data shows a very strong coefficient of determination of 0.9495 and 0.9994 for the minimum and maximum value, respectively. It thus implies that the exponent value of 1.42 is typical for any English text write-up. Since we are particularly interested in engineering write-ups, it seems that the exponent value of 1.42 may also be taken as a representative of this type of documents.

Conclusions

The objective of this paper is to report on the correlation properties in an English scientific text by means of a random walk. This has been achieved by showing how the proposed methodology can be used to judge the quality of engineering write-ups. In particular, the relationship between expression (1) and the quantitative level of the quality of the write-up as well as the results obtained in this work is explained. For example, how D and T affect the write-up quality and why the large exponent value of 1.4 can suggest a very good write-up are clearly stated. It is shown that for a random walk propelled in space using the pseudo-random number generator, the empirical relationship between the average distance and the time step is $D_{av} \propto T^{0.5}$, while for a random walk propelled in space using an engineering write-up this empirical relationship on the average becomes $D_{av} \propto T^{1.42}$. Therefore any engineering write-up may be likened to a form of random process characterised with an average exponent value of 1.42 to two decimal. It can be suggested that a very good engineering write-up will have an estimated exponent value very close to 1.42. This present study confirms that the average characterising exponent estimate of 1.4 is much greater than 0.5 (literature result) for a pure random motion. The exponent of 1.4 ensures a longer average distance covered in a very short time lapse compared to the exponent of 0.5.

Acknowledgements

The authors gratefully acknowledge the Postgraduate School, University of Ibadan, for providing information utilised in testing the theory proposed in the current work. In particular, thanks to Professor A.I. Olayinka and his research team. The words of encouragement of Professor P.O. Aiyedun on this research work are also acknowledged.

References

1. E. Petridou and K. Sarri, “Evaluation research in business schools: students’ rating myth”, *International Journal of Educational Management*, **2004**, *18*, 152-159.
2. A. I. Olayinka, L. Popoola, M. Fabunmi, and O. Fagbola (Ed.), “Synopses of Abstracts, Ph.D. Theses 2000-2002”, Postgraduate School, University of Ibadan, Ibadan, Nigeria, **2002**.
3. S. B. Lee and H. Nakanishi, “Markov matrix analysis of random walks in disordered continuous media”, *Physica A: Statistical Mechanics and Its Applications*, **1999**, *269*, 378-394.
4. Z. Jia, J. You, R. Rao, and M. Li, “Random walk search in unstructured P2P”, *Journal of Systems Engineering and Electronics*, **2006**, *17*, 648-653.
5. Y. X. Zhang, J. P. Sang, X. W. Zou, and Z. Z. Jin, “Random walk on psercolation under an external field”, *Physica A: Statistical Mechanics and Its Applications*, **2005**, *350*, 163-172.

6. C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, M. Simons, and H. E. Stanley, "Finite-size effects on long-range correlations: Implications for analyzing DNA sequences", *Physical Review E*, **1993**, *47*, 3730-3733.
7. S. V. Buldyrev, A. L. Goldberger, S. Havlin, C-K. Peng, M. Simons, and H. E. Stanley, "Generalized Lévy-walk model for DNA nucleotide sequences" *Physical Review E*, **1993**, *47*, 4514-4523.
8. P. Kokol, V. Podgorelec, M. Zorman, T. Kokol, and T. Njivar, "Computer and natural language texts - a comparison based on long-range correlations", *Journal of the American Society for Information Science*, **1999**, *50*, 1295-1301.
9. A. Schinner, "The Voynich manuscript: Evidence of the hoax hypothesis", *Cryptologia*, **2007**, *31*, 95-107.
10. W. den Hollander and P. W. Kasteleyn, "Random walks with 'spontaneous emission' on lattices with periodically distributed imperfect traps", *Physica A: Statistical and Theoretical Physics*, **1982**, *112*, 523-543.
11. G. H. Weiss, I. Dayan, S. Havlin, J. E. Kiefer, H. Larralde, H. E. Stanley, and P. Trunfio, "Some recent variations on the expected number of distinct sites visited by an n-step random walk", *Physica A: Statistical and Theoretical Physics*, **1992**, *191*, 479-490.
12. J. Feder, "*Fractals*", Plenum Press, New York, **1988**.

Appendix

52 Popular Alphabets found in the text message.
abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ
16 Agric. Engrg. Case4: Prof. Lucas "Design fabrication and evaluation of a beniseed (Sesamum indicum L.) oil expeller"
Some physical and mechanical properties of two benis
eed accessions were determined at five moisture leve
ls and used in the design of beniseed oil expeller.
A portable expeller for beniseed was fabricated base
d on the results of the determined properties. vario
us engineering values for the oil expeller has avera
ge capacity of 10kg beniseed per hour which, were st
ated. The efficiency of the expeller in terms of oil
recovery from the seeds as affected by wormshaft sp
eed and seed moisture content was reported for the t
wo accessions of the beniseed investigated. The maxi
mum filtered oil recovery for the two accessions was
observed at wormshaft speed of 45rpm and for seed m
oisture content of 5.3%. The estimated cost of the d
esigned oil expeller was stated with a clause that m
ass production would reduce the cost drastically.*
22 Mech. Engrg. Case3: Alabi Prof. "Fractal analysis and characterization of tread patterns of automobile tyres"
The author considered the fractal dimensions for the
tread patterns of selected 22 new and used tyres (1
2 brands). Preliminary studies on some familiar and
non-familiar objects having arbitrary shapes were de
veloped using fractal trees and the development of s
elf-similar fractal curves through the process of sc
aling and line segment substitution. Hence, the stud
y considered the two-dimensional (2-D) and three-dim
ensional (3-D) space fractal analyses of the tread p
atterns on new and used tyres. From the preliminary
studies carried out, the fractal dimensions obtained
in all cases considered were mostly identical with
the values quoted in the literature; lying between 1
.0 and 3.0. Similarly, 2-D space fractal dimensions
obtained in all cases for the printed tread patterns
were found to lie between 1.0 and 2.0. Moreover, 3-
D space fractal dimension of the tread patterns cons
idered, approached asymptotically the value of 3.0 w
ith increasing tread depth. Thus based on the values
of the fractal dimensions obtained, it is possible
to characterize the degrees of wear of the tyre' tre
ad pattern.*
20 Food Tech. Case2:Akingala Prof. "Functionality of flour in relation to physical and chemical properties of seeds of selected cowpea (Vigna unguiculata) varieties"
The functionality of flour in terms of Brdender visc

oamylograph pasting properties and thermal propertie
s, in relation to seed chemical composition and phys
ical properties, and product sensory quality was det
ermined for 28 cowpea varieties. Significant differe
nces were noticed in paste viscosities with hot past
e viscosity properties. Presoaking of seeds for 4 ho
urs before dehulling produced optimum particle size
distribution and highest HTPV in cowpea flour. Therm
al properties were significantly different with tran
sition enthalpy change (#H) explaining 80% while T-P
eak and T-onset explaining 14% and 5% of the variati
on in flour thermal properties respectively. Starch
and protein are the dominant macromolecular variable
s determining functionality of cowpea flour. The aut
hor suggested HTPV and #H as the important indices f
or flour production. There was a negative influence
of protein on HTPV an #H. This according to the auth
or might be due to reduction in starch concentration
.*
16 Agric. Engrg Case1: Igbeka Prof. Tillage systems, water management and their effects on soil properties and vegetable production.
Th author investigated the effect of no-tillage (NT)
, slashing (SH), ploughing (PHO), ploughing plus bed
ding(PHB) and two planting techniques; broadcasting
(BR) and drilling (DR) on soil physical and chemical
 properties using Randomized Complete Block design (
RCB). The saturated hydraulic conductivity in the 0
- 10cm layer was in the order PHO>PHA>PHB>NT>SH and
were significantly different at 5% probability. Soil
moisture potential at 10cm depth was least in NT(-48
cb) while PHA had the highest value (-32cb). At 0-10
cm, there were no significant differences in chemica
l properties while at 10-20cm depth, PHA was most ef
fective in enhancing soil nutrient status. Broadcast
ing had the best crop growth with PHB. watering was
recommended three times a week at 83.2% relative hum
idity and average rainfall 15.9mm*

Note that 16 on line three (3) indicates that the Abstract of the work supervised by Prof. Lucas was read from a document formatted into 16 lines, and the * on the last line (line 20) indicates the end of the Abstract. Thus 16 and * are format requirements to enable the computer to determine the beginning and end of each of the Abstract studied.