

Zipf's law in hierarchically ordered open system

Tadeusz Zipser · Magdalena Mlek ·
Wawrzyniec Zipser

Accepted: 29 June 2011 / Published online: 26 July 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract It was observed that inserting the arbitrary concentration into the simulation process makes the rank–size curves more similar to Zipf's law graphs than simulations based on initially equal size of each node. The hierarchical network's chain hypothesis based on the assumption that the prior cities' size differentiation, resulted from geographical determinants or historical events can constitute a set of priorities affecting the contacts. "Hierarchical chain" with different influence in the "up" direction as opposed to the "down" direction in hierarchy can be compared to the structure of sentences in a language, as well as to functioning of economic initiatives.

This setup has one important property: it allows for a part of the system to be cut off without affecting organization of the system. When an element is deleted or added, there is no evidence, which would indicate that the system was changed. Considering the easy 'truncation' or 'expansion' properties gives the proposed system high flexibility. Systems linked by such hierarchy lead to a clear distribution following the Zipf's law, if they are "open" i.e. directing most of the contacts outside of the system, requiring handling for only 'flow-through' contacts.

Special attention was devoted to an assessment of stability and endurance towards disruptions in ideal hierarchical order, so as to examine at which point the disturbance disrupts the Zipf's law.

The comparison of these modelings with recently accomplished selftraining modeling of settlement system reveals important relations between the parameter values emerged in that experiment and the hierarchical network.

T. Zipser (✉) · M. Mlek · W. Zipser

Department of Spatial Planning, Wrocław University of Technology, Bolesława Prusa St. 53/55,
50–317 Wrocław, Poland

e-mail: tadeusz.zipser@pwr.wroc.pl

M. Mlek

e-mail: magdalena.mlek@pwr.wroc.pl

W. Zipser

e-mail: wawrzyniec.zipser@pwr.wroc.pl

Keywords Rank-size · Intervening opportunities · Simulation · Hierarchy of cities

JEL Classification R12

Das zipfsche Gesetz im hierarchisch geordneten offenen System

Zusammenfassung Man beobachtet dass die Einführung einer arbiträren Konzentration in den Simulationverlauf macht die Ranggrößenverteilung der Städte der Zipf-Regel ähnlicher als in Simulationen die nur mit gleichen Zonengrößen starten. Die Hypothese der hierarchischen netzartigen Kette bezieht sich auf die Tatsache dass auf Grunde der geographischen und geschichtlichen Determinanten eine ursprüngliche Vorzugsreihe der Städte im Kontaktenbereich entsteht. Solche hierarchische Anordnung mit differenzierter „nach oben“ und „nach unten“ Wirkung kann man mit Bildungsweise eines Satzes in der Sprache und Leistung einer ökonomischer Initiative vergleichen.

Diese Netzart ist darin wichtig dass sie erlaubt dem System im Falle des Schwundes sowie des Emporkommens eines Elements irgendwelche Störung zu vermeiden und selbst dann keine Änderung zu bemerken. Diese Eigenschaft versichert eine grosse Flexibilität. Nach diesem Vorbild entstehendes System eine klare, mit Zipf-Regel übereinstimmende Größenverteilung erlangt wenn es offen ist und weitreichende Kontakte ausübt, die eher blosser Vermittlungsbedienungen fordern.

Besondere Aufmerksamkeit wurde der Haltbarkeit und Widerstandsfähigkeit gegen Deformationen und Unregelmässigkeiten gewidmet. Demzufolge wurde es überprüft wie eingeführte Abweichungen von der idealisierten Hierarchiestruktur die Zipf-Regelartige Verteilung bedrohen können.

Der Vergleich dieser Modellierung mit neuerdings eingesetzter selbstgesteuerter Simulation des Siedlungssystems zeigt wichtige gegenseitige Verhältnisse im Parameterbereich.

1 Introduction

The growth of cities and the existence of very stable distributions of city sizes is still an actual and inspiring problem. A theoretical way to explain the main mechanism of observed regularities seems to be necessary and useful (Zipf 1949).

The research focused on tracing changes regarding rank–size curves' parameters and shapes, obtained as a result of simulation aiming at achieving the state of contacts equilibrium of the simulated system; the changes were related to the modifications of the initial loads arrangement. As a result of the theoretical and natural networks analyses, it was observed that inserting the arbitrary concentration into the simulation process makes the rank–size curves more similar to Zipf's law graphs than simulations based on initially equal size of each node. Moreover, natural networks simulations that used asymmetrically located concentration generated better results than modeling based on centrally positioned concentration (Zipser and Mlek 2005; Fig. 1).

The simulation was operated on seven types of theoretical, regular networks, characterized by central symmetry in several axes. In each network the simulation was

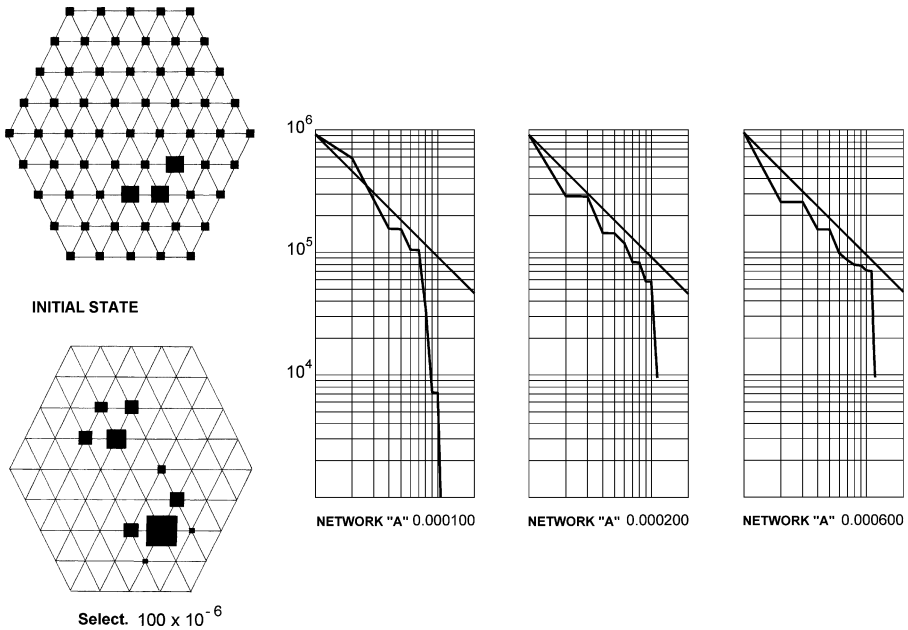


Fig. 1 The results of simulation with arbitrary involved asymmetrical concentration. Three selectivities

conducted with two types of initial loads arrangement: the first type based on initially equal size of each node, and the second took into account asymmetric concentration, where three neighboring nodes received increased loads.

Simulation results, analogous to the aforementioned ones, but based on natural settlement networks are achieved in two kinds of network—road and train network in each of the eight areas (Austro-Hungarian Monarchy (the end of XIX c.), France, Germany, Great Britain, Italy, Scandinavian Peninsula, Spain, and United States). The concentration was arbitrarily located either in the capital or in one of other important cities. Considering the common imperfection of natural networks three different methods were used in order to assess the location of the center of symmetry, i.e. to assess the level of asymmetry of the initial load concentrations' positions (Zipser and Mlek 2005; Fig. 2).

Now we present the hierarchical network's chain hypothesis based on the assumption that the prior cities' size differentiation, resulted from geographical determinants or accidental historical events, was strengthened constituting the system of preferences for distant contacts. The simulation was applied to the networks defined not by common spatial relations (based on distance), but by hierarchical order not dependent on space. By means of this kind of hierarchical order, the cities rank positions defined the sequence of potential destination penetration.

In the article a range of simulation variants was discussed, based on various modification types of the hierarchical network's chain, multiply initial loads displacements, the system size, the selectivity parameter value and other, treating them as a possible aberration of 'ideal' model hierarchy. While analyzing Zipf's law graphs obtained due to the simulation process a special attention was devoted to an assessment

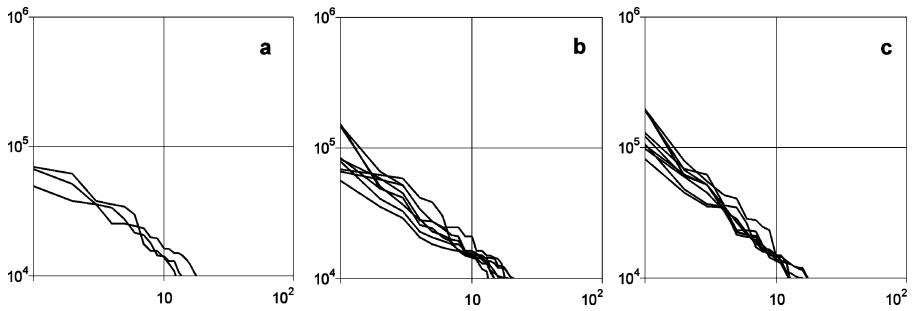


Fig. 2 Spanish rank–size plots for simulation variants conducted on road network and: (a) initial equal distribution of potentials; (b) initial distribution of potentials with concentration situated in Madrid; (c) initial distribution of potentials with concentration situated in Barcelona

of stability and endurance towards various disruptions in ideal hierarchical order, so as to examine at which point the disturbance of hierarchical structure disrupts the system characteristics consistent with Zipf’s law.

The work indicated two ways of investigating the sources of Zipf’s law regularity. Both arbitrarily inserted disruption in the simulated systems taking the shape of a significant, asymmetrically located concentration of origin and destination loads, and substitution of spatial nodes accessibility by its hierarchical ordering of special type—resistant of amputations and intrusions, caused the rank–size curves to become similar or nearly identical to the ideal Zipf’s law graph. It seems that it is possible to interpret both sources of Zipf’s law regularity as compatible based on a defined events sequence or even its mutual intensification.

2 Hierarchical network chain hypothesis

A hypothetical sequence could look as follows. We have an offset strong ‘mass’ appear within a network structure of regular communication links, i.e. the ‘mass’ being a collection of contact origins and destinations arising out of external determinants or arbitrary decisions.

This creates, as part of the process of establishing contacts and linked concentrations, a distribution of system units, which starts to follow the Zipf’s law.

Such notable diversity establishes contact preferences based on efficiency rather than minimum distance, thus introducing a hierarchical scheme of contact priorities.

This hierarchical network strengthens, smoothes out and supplements the rank/size distribution following the Zipf’s law, and it further supports this distribution as new units appear, increasing number of origins and destinations in the system. This happens when the system is open i.e. directing most of the contacts outside (of the system).

As presented later in the paper, the concept of the mechanism generating the Zipf’s distribution separates from the elements’ spatial/distance determinants, as used for example as the basis for Lösch (1944) spatial theory. If we consider the elements of interest to be cities, then in view of the above, the contacts will not rely on a standard

transport network but rather a hierarchical communications (i.e. “communicate” vs. “commute”) network.

Another, more general, stipulation is that both communication and contact responses must be fairly simple, otherwise the occurrence and persistence of the universal order/size rule would be uncommon and inexplicable, especially in view of the many different areas, where it proves correct.

Nevertheless, in each of the mentioned areas, there are multiple factors which affect the given values and occurrences—which explains the imperfections observed when comparing the real sets with the expected distributions. What is important, the observed deviations ‘oscillate’ according to a given scheme, giving us the steady state of the population of the given city.

This suggests an attempt to replace the flow system based on a transport infrastructure with a hierarchical flow system, which does not take into account the geodesic spatial distribution of points.

3 Hierarchical network contact flow algorithm

Concluding the earlier stipulations lead to a proposal of a modeling algorithm used to verify the generation of Zipf's distribution. The initial conditions assume specific relations between the cities arranged in a hierarchical order—i.e. where each city has an established relation with the city ‘above’ and with itself. This means a contact priority with the first city, then the second . . . then the third . . . and so on. This hierarchical chain terminates within the own zone (city) and in the modeled simulations the generated contacts and their balance reflect the above destination penetration sequence in view of the contact terminations originating from a given zone (Fig. 3).

The remaining unfulfilled contacts will be directed to lower zones in this hierarchy, however without assuming any specific order. This is carried out by proportional (based on zone size) division of all available selectable destinations. This setup has one important property:

It allows for a part of the system to be cut off without affecting organization of the remaining part of the system. I.e. cutting of a city or several cities at the bottom of the hierarchy, as well as any from the top (including the largest city) will obviously just cause the disappearance of the relevant relations. When they are deleted, there is no evidence, which would indicate that the system was changed in any way. Same thing pertains when a city is dropped from the middle of the hierarchy. Similarly, adding a city at any place of the series only requires definition of several input and output relations, but does not require change of the existing ones. As opposed to the above, in cases of stacked (including Christallerian) hierarchical systems, when a city is deleted it either leads to imbalance caused by weakening of one of the subordinate groups—or even worse, total disintegration of the system if an important intermediate link is lost. Considering the easy ‘truncation’ or ‘expansion’ properties gives the proposed hierarchical system high flexibility and reflects the colonization and development conditions of the set of cities which comprise such a system. This also explains the fairly quick adaptation of this set to such extreme changes as moved borders, which could be observed among others based on Poland after World War I and World War II.

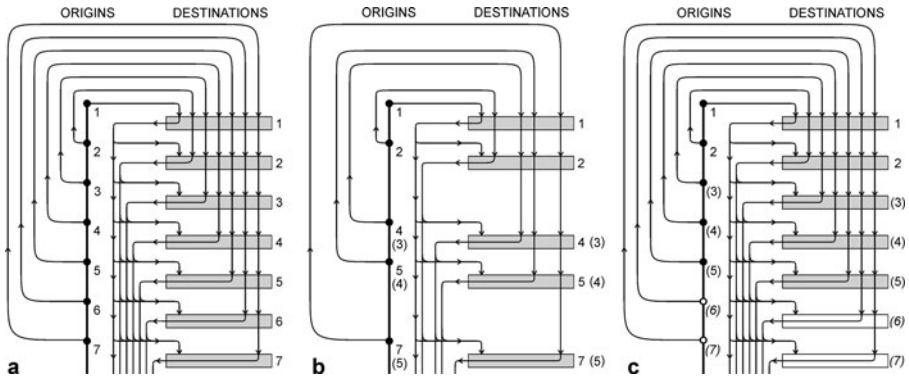


Fig. 3 Hierarchical network chain scheme. Penetration sequence: (a) version without own zone priority (b) structure after dropping zones 3 and 6 (unchanged penetration sequence) (c) new structure after dropping zones 3 and 6

It should be noted that a similar rule also applies to language. People with quite different word bases, or even with quite low knowledge of a given language, are still able to communicate and formulate a large number of comprehensible sentences. Similarly, when certain words are dropped, the text can be reformulated to maintain the original sense.

Applying the above rules, we thus build a system, where the set of unitary relations is not strictly defined, even though we have a clearly defined total number of people as well as the number of cities comprising this system. This is caused by the unknown distribution of people among the cities in the system. We therefore are able to assign various number of people to sites within the system, for which we only have the relation scheme towards the rest of the system.

Starting with the definite distribution of destinations determined suitably to a given task (e.g. place of work, anticipated clients of services, etc.) and the hierarchical table of mutual accessibility of zones, calculations of movement exchange are made by means of the intervening opportunities model (CATS 1960). The *p* value or selectivity is, of course, determined at the beginning of calculations. This selectivity may be constant during calculations or changed in successive approximations.

After the calculation of movement exchange all surpluses of acceptations of destinations ascribed to zones and all insufficiencies of destination acceptance in other zones are obtained on the basis of compilations prepared according to zones (“balance conditions”). Moving to the next approximation, the “shifting of goals” is carried out, i.e. surplus zones are now ascribed such a number of destinations anticipating acceptance as the number of acceptances recorded in the preceding calculations (Zipser 1990).

This type of model structure, which is used in the general shift formula, can also be presented in the form of formulas presented below, which partially take into account the influence of the hierarchical network. I.e. the number of contacts terminated at zone *j* comprises of contacts *G_j* received from zones lower in hierarchy and contact *D_j* from higher zones.

$$B_j = G_j + D_j$$

However with G_j also including contacts from own zone (when $i = j$). Therefore:

$$G_{j \neq 1} = \sum_{i=j}^n [e^{-pa_f} - e^{-p(a_f+a_j)}] Z_i$$

where: p —selectivity parameter, density of probability of destination acceptance; a_j —number of destinations in zone j ; a_f —number of destinations considered before a_j (included in zones higher in hierarchy than j), thus:

$$a_f = \sum_{l=1}^{j-1} a_l$$

Z_i —number of origins in zone i , Z —number of all origins in the system.

For zone 1 located at the top of the hierarchy we can therefore have:

$$G_1 = \sum_{i=1}^n [1 - e^{-pa_1}] Z_i = [1 - e^{-pa_1}] Z$$

With the second part defined as:

$$D_{j \neq 1} = \sum_{k=1}^{j-1} [e^{-pa_f} - e^{-pA_k}] Z_k \frac{a_j}{A_k}$$

$$A_k = \sum_{l=k+1}^n a_l \quad \text{when } a_f = \sum_{l=1}^k a_l$$

Of course this will look differently if the own zone will be placed in the same hierarchy as zone 1. However, this case as well as other disturbances of the model hierarchical network are sufficiently controlled in the simulation model by appropriate network records. The above formulas serve only as illustration of the essence of the “clean” experiment, rather than a computational tool. As one can see, the position of zone 1 is special because the inflow of contacts is independent of the origin and destination distribution between the zones. It will keep on rising until the value of the exponential cumulative distribution multiplied by the number of contacts in the system becomes equal to the number of destinations at the given zone.

The special situation in our model arises from the fact that we are not thinking about direct migration of contact locations, but rather diversification of the origin and destination masses within the system, which reflect the flow of contacts, organizing the long-term social/economical activities and the city’s position in a much wider area, where most of the contacts originating within the area are yet to be terminated. In view of the above, the modeled element is not only a origin or final destination, but rather a ‘middleman’ in long and widespread chains of action. This reflects the current economical and cultural mechanisms of civilizations in the XXth as well as major portions of the XIXth Century.

Considering this, each subsequent acceptance distribution in the cities of our system leads to distribution of origins and destinations within the system, in such a way that the size of the city allows for receipt and transfer of relations essential to the existence of the entire system. This is carried out by normalizing the number of origins and destinations after each iteration. The modeled flow gives and outlook on the load diversification for the different zones—in our case individual cities (which identify the zones).

It is thus expected that the repeating “waves” of contacts will subsequently lead to a certain balance, where the number of destinations in the zones will reflect a distribution according to Zipf’s law.

In comparison to normal “migration models” reflecting spatial concentration distributions, we introduce one fairly important modification. The number of terminated contacts (number of acceptances located in the destination zones) is not synonymous with the number of migrations. The acceptance distribution in this network simulation gives us an outlook on a certain, irregular need to serve these contacts, which should be reflected by the normal distribution of people. Considering this, we get an image of forces causing migration and distribution of attracting factors (mainly workplaces) rather than the actual migration of people. If this was not the case, we would either have to include all contacts within the set defined by the network or allow for a decrease of the number of inhabitants, caused by migration outside of the system, where such decrease would rise with “sharper” selectivity factors (i.e. lower p values).

Based on the above, we can formulate the following stipulations—hypothesis: the city size distribution following the Zipf’s law occurs in highly open settlement systems. Similar statement was done by Jokiel and Kostrubiec (1976) and Haag and Max (1995).

This is an important hypothesis, because many of the earlier interpretation attempts of the Zipf’s law, tried to tie-in it’s existence to a system of cities, assuming almost automatically a rather closed economy or autarky of such a system. Such characteristics can be presumed looking at Christaller’s Central Place Theory, which it would be rather difficult to make comply with the Zipf’s law.

In order to verify the said hypothesis or rather the general cumulative hypothesis assuming their combined effect, we have performed many models using a discrete system.

The quantitative characteristics of the modeled system are as follows:

1. *Number of zones.* We assumed sets with 20 and 30 zones. As the result of application of principles assumed at the later stage of modeling, we had to introduce an additional zone, thus the resulting totals were 21 and 31 zones.
2. *Number of origins and destinations in the system.* For the set of 20 zones the number of destinations was assumed at a total of 1,000,000, for 30 zones it was the same (although with different distribution per zone).

With the introduction of the additional zone as the unit representing the surrounding area, it was assigned with 1,000,000 destinations and in one of the cases with 50,000,000 destinations.

In all model versions and at each stage of modeling we assumed that in each iteration the number of origins is equal to the number of destinations in each of the zones.

3. *Selectivity parameter.* The largest variation was applied to the p parameter defining the probability density in the negative exponential distribution, which is the key parameter defining contact distribution. Its value depends on the number of destinations, or in other words, on the scaling of elements of the contact “absorbing mass”.

We thus used values from $p = 0,0001$ down to $p = 0,00000005$, with most of them oscillating closer to the later value.

In each variant, we would either have one parameter assigned to all of the zones or a set of different parameters for each of the origin zones. What is important, the selectivity parameter assigned to the origin zone remains constant during penetration of subsequent zones, where each of them can have its own selectivity parameter. This selectivity, which is synonymous with the acceptance probability of a randomly chosen destination, is a specific characteristic of the contact rather than a given attracting factor for this contact.

4. *Computation series.* Each modeling variant included a series of iteration, where each subsequent run used as the input the situation obtained from the previous iteration, without any other external input. The number of iterations varied from a few runs when checking a given data set to even a few hundred (250 or even 300) when trying to verify the stability of the generated configuration (Fig. 4).

We have created 89 different models, where for each iteration the results for each of the zones were plotted in the form of superimposed lines, visualizing the changes occurred throughout the modeling sequence. The plots were observed on the screen in natural scale, but each iteration for each modeling was also illustrated by a double logarithmic scale plot, which allowed for comparison with the plot shape under Zipf's law.

In total we thus viewed 89 plots of the first type and around 6700 plots of the second type. Selected plots are included in this paper to illustrate the results, relevant to the interpretation of regularity and evaluation of the main hypothesis.

The large variety of investigated models arises out of two reasons. The first reason was to pick appropriate selectivity parameters. The second, more important reason was to evaluate different input conditions as well as the different hierarchical network configurations. Whilst the first reason relates to the need to ‘calibrate’ to the best extent the applied mechanisms, the second one is used to check the sensitivity of this mechanism to interferences, which naturally will occur and which are unavoidable. An ideal model and most importantly its main part—the hierarchical network structure, is unlikely to expose a clear/sterile operation in the complex reality of a real region or country. Considering this, it is necessary to check the possible extent of deformations of the model setup, which will still allow for a relatively low impact onto operation purity, or a change of effects comparable to real world observations of settlement systems.

Considering this, most of the modeled variants considered changes in the data sets, either in (a) reciprocal tables of zone availability—i.e. change of the hierarchical network itself, (b) tables of initial distribution of origins and destinations, along with the influence of the various concentrating factors, and the earlier mentioned (c) different values of the p (selectivity) parameter, used either for all of the contacts or predefined individually for the outgoing contacts for each of the zones.

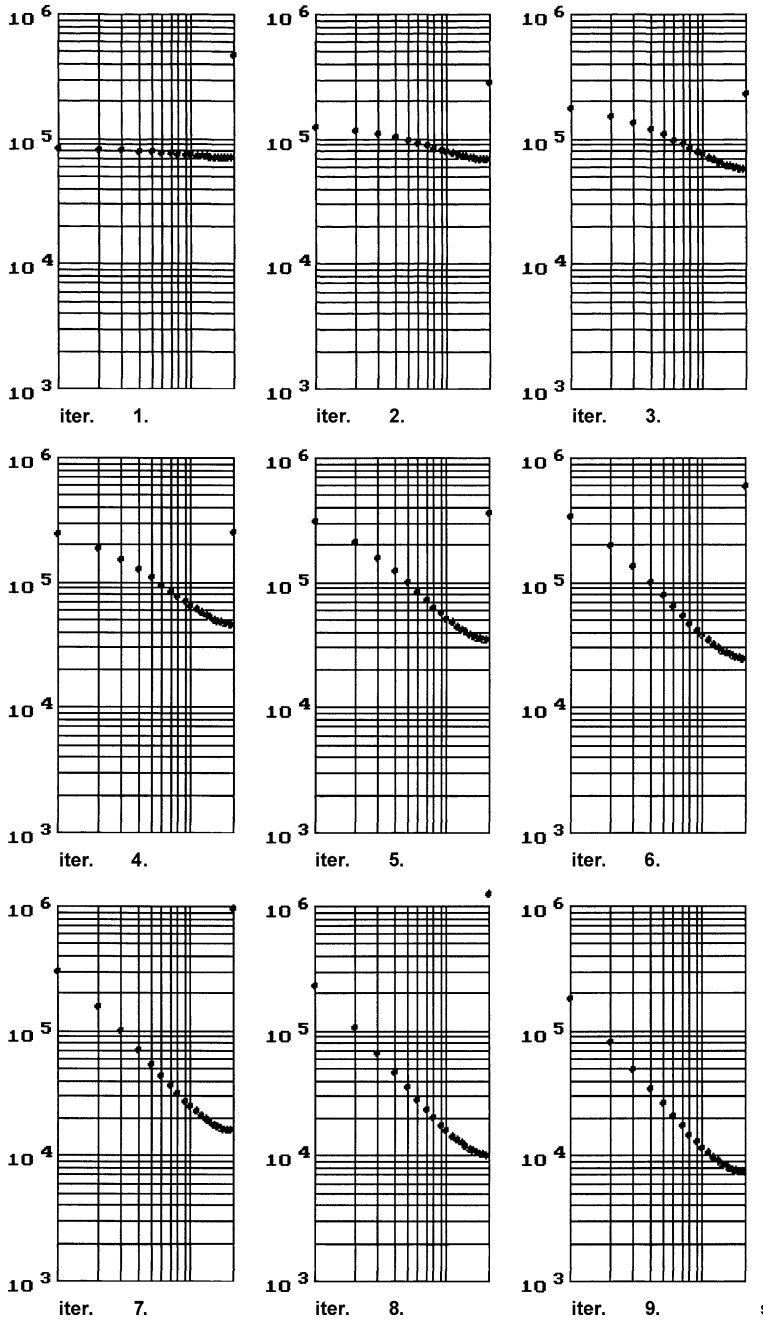


Fig. 4 Sample order/size plots for the subsequent iterations of the modeled setup, taking into account the unit representing the last surrounding of the modeled system

4 Modeling

4.1 Building regularity

The objective of the first series of variants operating using the hierarchical network (“m 1”) was to identify the selectivity interval, which will allow the results to match most closely the Zipf's law distribution. Values above 0,0000001 resulted in double logarithmic plot at an angle of around 135° and straight characteristic, except for the right end pertaining to the zones lowest in hierarchy.

Because the initial conditions were standard, i.e. all zones had the same number of origins and destinations equal to 50,000, this meant that higher growth is linked to higher position in hierarchy or lower ordinal number of the given zone. Thus after just few iterations—in run 5, 6 or 7 we already start to observe the characteristic Zipf's law properties along with the earlier mentioned disturbance—i.e. flattening of the plot for the few (starting at around 6th) lowest numbers—making the plot shape exhibit a “hockey stick” distribution (this is how we will refer to this deviation shape from an ideal plot).

The next series applied p values lower than 0,0000001 (“start 11”). In iteration 30 we start to observe the proper angle (135°) without dominants, whilst subsequent iterations bring on a slow regular increase of the slope, thus we can confirm strong stability of the Zipf's distribution, except for the also evident here “hockey stick” pattern).

4.2 Disturbances on the hierarchical chain network

In the subsequent step, we introduced modifications to the distance table. The change for each origin zone consists of assuming in the first table row (distance range) the zone itself rather than zone 1.

Results with a selectivity $p = 0,0000001$ are similar to “start 11” however p with a slight sinusoidal distortion—a small shortage in zone 1 at the top of the plot. Continuing the modifications, for the first distance range we now assume zone 1 along with the origin zone itself, i.e. returning somewhat to the original setup. For $p = 0,0000001$ we obtain very good slope at iterations 41 to 47, however with a significant flattening towards the end.

Further changes to the accessibility table introduce more significant disturbances to the hierarchy structure. Taking the original setup “m 1” we have now created several “shelves”, where we assumed several (2, 3 or 4) zones within the same distance range, as opposed to the earlier separate range setup. This means lack of hierarchical differentiation—synonymous with same destination acceptance range for the given hierarchical distance. Such ‘shelves’ are basically placed in every second row as well as one per each row (“mplat”).

Again, the best results here were obtained with $p = 0,0000001$, i.e. achieving proper slope with slight domination of zones 1 and 2 and an increased “hockey stick” effect.

Comparison of the two distortions: the “mplat” network and chaotic set of zone selectivity parameters (random distribution of values from 0,000001 to 0,000200)

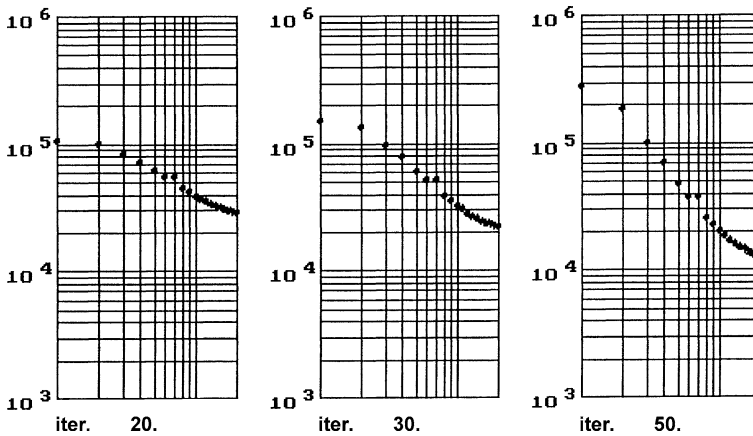


Fig. 5 Modeling with a modified hierarchical network—three patron zones with subordinate parts of the system

showed that the resulting hierarchical order was maintained, however with an increased slope and quicker concentration.

Implementing further hierarchy disturbances:

We introduced changes to the distance table (“m plat 2”) so that “shelves” are present in each of the rows—the characteristic and slope still does not significantly diverge from the original plot, however small shelves and set-off remain visible with a slight sinusoidal profile.

4.3 Split of the hierarchy apex

This disturbance has a deeper effect, as in place of zone 1—i.e. the first field of the table (“mpatr”), we use every third row of zone 2, 4 or 7. This is as if we had three patron zones with subordinate parts of the system (Fig. 5).

As well as (“m three 1”) consisting of introduction of three independent threads at the beginning of the table (for zones 6 and 12—inclusive). Which bring the relevant zones 1, 2 and 3 to the top of the hierarchy. Fourteen or eight last zones have rows built normally, whilst the first range includes all three patron zones at once (Fig. 6).

4.4 Uneven starting masses

Here we assume an arbitrary change of the masses—i.e. the number of initial origins and destinations.

First we triple the mass for zone 1 (150,000 destinations and origins instead of the initial 50,000), and in other variants we apply the same change to zones 6 and 13, whilst also decreasing the mass in last four zones to one half of the initial value.

The model exhibited a strong domination of zone 1, when it was increased, however when the same was done to zones 6 and then 13, in both cases we achieved proper slope (around 135°), where the oversized zones were gradually reduced. This

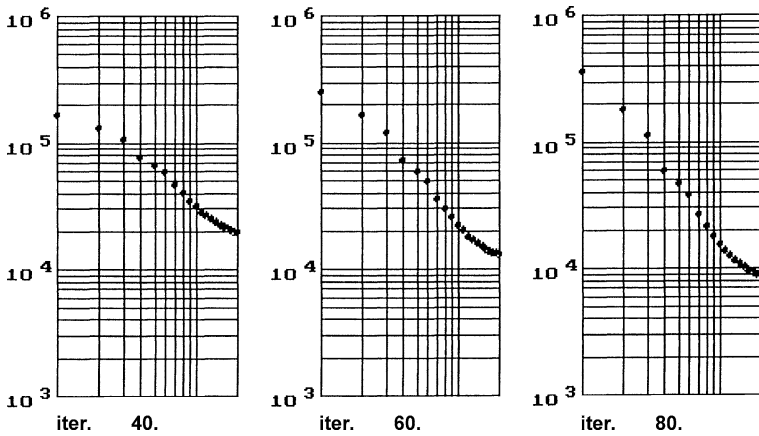


Fig. 6 Modeling with a modified hierarchical network—three independent threads at the beginning of the table

profile proved stable, however the initial jump for the privileged zone remained visible (one should remain that the change of places on the axis, plotting the values by size, would only exhibit a slight disturbance).

A reverse approach—reduction to one quarter the initial mass for zone 1, emerged a proper slope excepting delayed growing of zone 1.

4.5 Chaotic modification of hierarchical relations

The next step consisted of introduction of chaotic modifications, distorting the theoretical regular structure of the assumed hierarchical network. Random swapping zones among almost all rows and sometimes locating even a few of zones in improper row (in total up to around 30 positions), however without breaching the increasing differentiation rule (i.e. decreasing common ranges), caused insignificant distortion of the plot, which had good slope, high stability and almost always a slight “hockey stick” flattening towards the end.

Introduction of further modifications, consisting of zone shifting, flattening (increasing common ranges—i.e. removing distance differentiation) as well as several separations (splits) of the normally common ranges, lead to an increased number of splits, local protuberances etc. however without affecting the general plot characteristics, especially if one moves the points to the proper positions ordered by size.

4.6 Increased number of zones

The next modeling run was performed for a 1,5 times larger set of points (30 zones). Selectivity $p = 0,0000001$ and regular hierarchical network “me 30” gave about 40 iterations with a very regular Zipf's distribution with a slight “hockey stick” on the end. With further iterations we obtained an increased slope, but with sharper selectivity $p = 0,00000005$ we have an almost ideal profile and stability (Fig. 7).

Arbitrary reduction of the initial masses of the last four zones in hierarchy (zones 27, 28 and 29 with 3300 destinations and zone 30 with 1300 destinations) almost

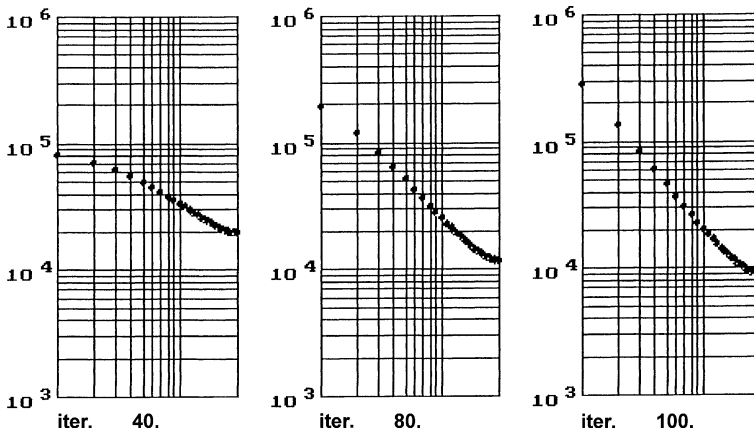


Fig. 7 Modeling of the hierarchical network for a set of 30 zones with selectivity 0,00000005

totally reduced the “hockey stick” end, causing the end plot results to drop for the last zone (similar to what is observed in real world results).

In another variant, a tripled zone 5 was effectively reduced.

5 Zipf’s law for cities and lexical analogies

One of the fields where Zipf’s law is especially often and clearly referenced is linguistics. Zipf’s law, which here is called Estoup–Zipf’s law, pertains to word frequency in text. It defines that the product of the rank (subsequent number on the frequency list) of a given word and the number of its occurrences (in different grammatical forms) in text is a constant, relative to the text length. This property validates quite well, regardless of the text type (anywhere from fiction to scientific papers) or language (Estoup 1916).

Comparing the plot with the list of words to which it pertains, one can easily see, that the more often occurring words, are more general and vague whilst serving more organizational purpose in a given sentence. This is obviously why such words are used more often, as they are more fit or even absolutely vital in certain situations. Words from the other end of the plot pertain to much more specific/narrow subjects, carrying much more ‘valuable’ information and therefore there is much lower probability of their occurrence at any place in the text.

Generally, it can be summarized that in order to create a sensible message, one has to use words from different places of the rank list i.e. from different parts of the plot. Reversing this—the different frequency of occurrence of words is attributed to their different purpose in constructing meaningful sentences.

The presented hypothesis pertaining to cities and the regularity following the Zipf’s law applies a certain analogy of the linguistics phenomena. We assume that an effective economic or social activity is the equivalent of an important sentence in text. Such activity must also use certain substantial resources (raw materials, narrow type of production, certain group of people, specific qualities of the given area etc.) as

well as more general, yet still vital ones (financing sources, investors, import/export organizations, transport equipment, administration institutions, insurance companies, sources of energy etc.)

Summarizing: after the industrial revolution, especially in view of the current development of transportation or in general communication, as well as globalized economy, politics and culture, it is normal for different specialized subjects to interact and cooperate in shaping the economic/social activities. A diversified range of versatility and unique offers establishes a hierarchical scheme of potential relationships between these subjects. If a major percentage of such relations must be directed to a very broad range of potential destinations, to assure a sufficient probability for establishing a link, this means purely large diversification and thus integration of the civilization system.

6 Zipf's law vs. duality function of the city and its active population

Diversified contact frequency, which always arises out of certain activities, must also involve a similarly diversified (i.e. relevant to the distribution) 'personnel'. This explains, in view of the Zipf's law the interrelations pertaining to city size. This reasoning can or perhaps should be applied only to a certain group of employed city population, which can be determined empirically based on the location of the point on the plot (double logarithmic scale), at which the straight line suddenly bends. The influence of the rule is limited to just a part of the entire population of the cities. The aggregation on the plot (on the city size axis) of values, which form the upper right-angled triangle and the values, which create the 'base' at the point of the bend, can be interpreted as the product of two values (the sum of logarithms of numbers is equal to the logarithm of their product). Of these two values, one is a constant ('base') and the other is a variable (triangle). One can thus suppose that a certain group of people, lets call them 'system-forming', is responsible for the city size diversification in the system. These people are supported by another group, which is always proportional to the first one, at a certain constant ratio. In the cities where the 'system-forming' group is practically nonexistent, there is also not much justification for the existence of the supporting group, which consists of just residual (or comparatively 'seed') conglomerations. This could mean for example that in a million people city, around 2000 people are involved in shaping the above mentioned inter-city system cooperation of the first type, and for each such person there are around 500 other people responsible for handling internal city functions, as well as others, which are not participating at all in professional activities (especially because of their age). This would reflect a classical split of city population into the 'base' and 'service' sectors, or in other words exogenous and endogenous employment. Even if we assume a different ratio, as for example 1:1 for this 'system-forming' group, this would mean (subtracting about 50% people as not professionally involved) that for a 1 million population city, we are talking about only something like 250,000 people.

The "strong" version of Zipf's law assumes that there is a certain threshold limiting the lowest size of the last unit in the order/size distribution, relevant to the first unit size (Jones and Lewis 1990). With a limit of 500 people, we automatically obtain

an upper limit for the first city, and if we maintain the ideal distribution $P_j = P_1/j$ and the number of cities in the system equal to n we obtain:

$$P_1 = P_n \cdot n$$

and for a set (e.g. country) consisting of 800 cities (similar to Poland where we have 860 cities) we obtain $P_1 = 500 \times 800 = 400,000$ people. With the same lower threshold we could also assume that this pertains to the ‘system-forming’ people, and thus the real minimum required for proper functioning of the city must take into account the ratio of the ‘system-forming’ people, for example each of them ascribed to a minimum of 5 other people. This changes our calculations, where now we have the smallest city of 2500 people, which should be capable of existence within the set of other cities. With an 800 element set, this also defines the size of the first largest city at 2,000,000 people. This roughly reflects the situation in Poland, if we take into account that the Warsaw agglomeration is nearing 2,000,000 people, whilst the smallest Polish city (Wyśmierzyce) has 880 people. There are a total of 82 cities with population under 2500 people, and only 43 with population below 2000 people (Bagiński 1998).

Recently an other mode of simulation was developed in order to test the hypothesis concerning the influence of diversification of selectivity on the settlement system. In such case this variety of parameter values would be presumed a good indicator of needs, aspirations and real possibilities of local population.

Now we intend to achieve the compatibility of current and simulated sizes as good as possible. The aim is different than in previous experiments. As we already know precisely actual situation and all particular magnitudes we need the information about the process of modification of parameter in course of emerging final size. Such modification should be an individual characteristic of each zone.

It allows us to obtain precisely the actual size of every unit in the selected system of cooperating cities, without any diversification present initially. A necessary course of changes concerning the process parameters is recognized because the procedure is something like a self-training action. Further classification and interpretation of consecutive fluctuations and trend variations help to explain probable factors affecting reality.

The self-steering formula is fairly simple:

$$p_i = \frac{-\ln R}{B_i}$$

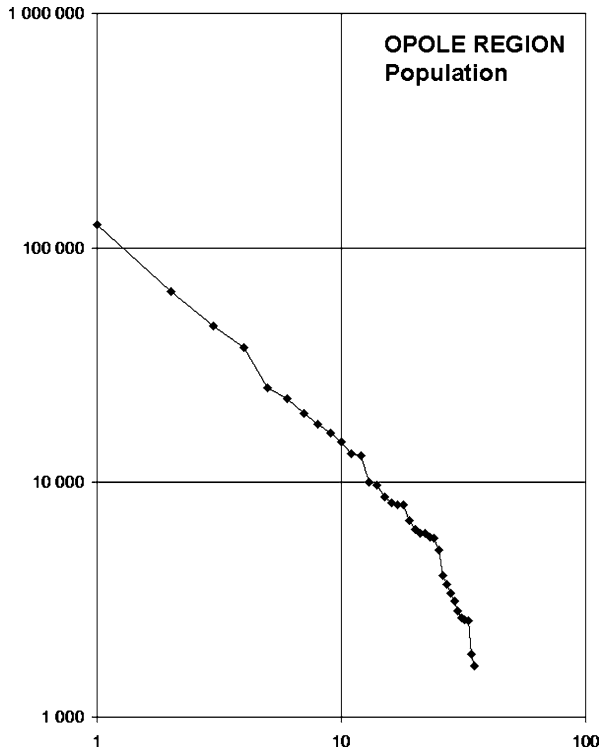
when the zone i is still too small, and

$$p_i = \frac{-\ln R}{A_i}$$

when that zone has to decrease to preserve its expected size,

where: R is an allowed fraction of contacts generated in zone, which are not satisfied by its own opportunities (usually $-\ln 0,001$ equal 6,908); B_i is the number of destinations (acceptations) received in given iteration; A_i is the expected number of destinations, and; p_i is the selectivity parameter for the next iteration.

Fig. 8 Opole voivodeship rank–size plot based on 2008 city sizes



To apply this second version of simulation—it means the selftraining algorithm we used the example of Opole region in Poland containing 34 cities. This region is one of the best representatives of rank—size rule as obeying very exactly the Zipf's law regularity (Fig. 8).

The replacing road network with the hierarchical (network) sequence of penetration brought very quickly the result of full and perfect accordance with real values. It was done starting with equal uniform population for all cities.

To achieve compatibility like that of 50 iteration there were insufficient to calculate 500 iterations in case of using the road network accessibility (Figs. 9 and 10).

The final set of modified selectivities (after 100 iterations) applied again for uniform distribution of population gave also very quick effect of full compatibility, what was never the case for road or geodetical networks.

It demonstrates that the Zipf's law regularity may be achieved not only in the way described above as the population of middlemen in very extent exchange of relations in an open system. It is possible to reach such regularity also with very moderate selectivities in a limited not large system, provided the values of this parameter are strongly diversified (until few dozen times) and the mutual contacts take place in hierarchical way. Perhaps such behavior is common now among the cities.

Fig. 9 Correlation plot of Opole voivodeship cities in the year 2008 (X axis), and their sizes gathered in 100 iteration of simulation based on hierarchical network carried with initial selectivity 0,0001 (Y axis)

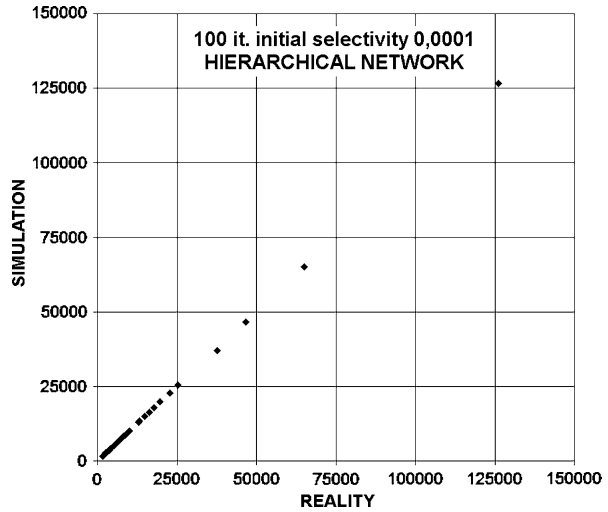
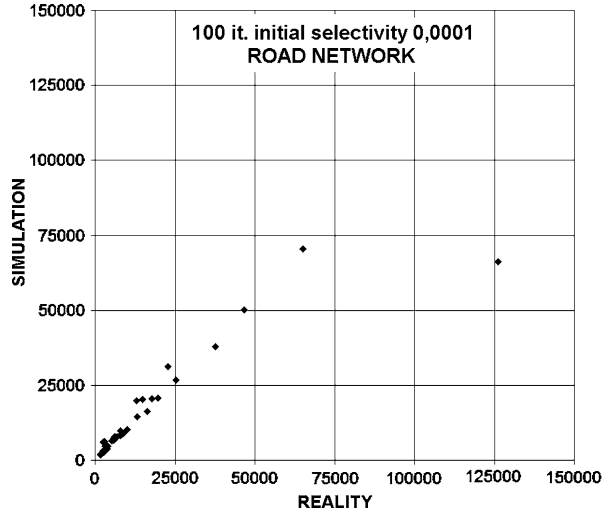


Fig. 10 Correlation plot of Opole voivodeship cities in the year 2008 (X axis), and their sizes gathered in 100 iteration of simulation based on road network carried with initial selectivity 0,0001 (Y axis)



7 Conclusions

We have performed three independent experiments with a common objective to locate balance states in contact exchange between zones, which can be attributed to cities of a region or country. In both cases the contacts were simulated using the destination and origin movement model, applying the “intervening opportunities” mechanism.

Both arbitrary modifications (“disturbances”) in the form of predefined significant concentration of origins and destinations at offset locations, as well as replacement of the spatial interavailability of zones by a linear ordered hierarchical equivalent, lead to a distribution of origin and destination concentrations similar or almost identical to an ideal distribution according to Zipf” law.

It therefore seems reasonable to propose the following interpretation of the observed distribution:

1. Asymmetric location of region's capitol can significantly affect (even with a highly developed regular communications network) the development of a specific order of city sizes.
2. It can constitute a set of preliminary priorities affecting the development of contacts important to development and existence of economic and social structures, when the physical distance barrier is no longer a factor. A very simple "hierarchical chain" with different influence in the "up" direction as opposed to the "down" direction in hierarchy can be compared to the structure of sentences in a language, as well as to functioning (creation and maintenance) of economic and social initiatives.
3. Even if initially starting with the same input masses, systems linked by such hierarchy lead to a clear distribution following the Zipf's law, if they are "open" i.e. directing most of the contacts outside of the system, requiring handling (certain number of people) for only 'flow-through' contacts.
4. The diversified population of cities in the hierarchical network results from the diversified number of terminated (or passed on) contacts, which need to be handled. Contacts should be widely interpreted as incoming or flowing-through 'goods' requiring further 'handling' (e.g. raw materials, semi-finished products, finished products—which have to be sold or exported, students—who have to be educated, graduates—who have to be employed, credits—which have to be obtained, loans—which have to be paid back, permits which have to be obtained etc.).
5. When it comes to Zipf's law distribution, these systems are fairly insusceptible to disturbances (splits, flattening or shuffling of the hierarchical chain, as well as disappearance/extinction of certain zones).
6. Provided, the hierarchical way of penetration is used by the set of units operating with strongly diversified selectivities, the Zipf's law regularity may emerge in limited, even closed systems of moderate size.
7. Considering that a combination of any number of Zipf's law systems still maintains a Zipf's law distribution (although certain 'shelves' can form) means that a system built by adding together "Zipf" subsystems will also be a Zipf's law system, and since the component subsystems are not identical, in the resulting system instead of clearly visible 'shelves' we only observe slight local disturbances of sections of the plot slope.
8. Independent formation and growth of the component subsystems explains the maintenance of Zipf's law distribution, even with significant modifications of the system rank positions. Creation of new and quantitative growth of old subsystems does not destroy the Zipf characteristic of the entire system, unless it itself destroys its local Zipf's law distribution.

In order to fully confirm this hypothesis, research should be performed in regards to real world coverage results, as well as for mutual contact conditions required for creation and maintenance of overall economical, cultural, educational, informative and administration/organization activities.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Bagiński E (1998) Małe miasta w strukturze osiedleńczej Polski. Wrocław, Tab. 1
- CATS (1960) Chicago area transportation study. Final report II
- Estoup JB (1916) Les gammes stenographiques. Institut Stenographique de France, Paris
- Haag G, Max H (1995) Rank-size distribution of settlement systems: a state attractor in urban growth. *Pap Reg Sci*, 74(3), 243–258
- Jokiel B, Kostrubiec B (1976) An attempt to determine the size of an open settlement system using the rank–size rule. In: *Economic models in regional development and planning*. Warszawa, pp 165–173
- Jones G, Lewis BD (1990) The four basic properties of rank–size hierarchical relationships: their characteristics and interrelationships. *Pap Reg Sci*, 68, 83–95
- Lösch A (1944) *Die räumliche Ordnung der Wirtschaft*. 3rd edn. Stuttgart 1962
- Zipf GK (1949) *Human behavior and the principle of least effort*. Addison-Wesley, Cambridge
- Zipser T (1990) A simulation model of formation of the settlement structure. In: *The model of intervening opportunities in theory and practice of territorial arrangement*. Wydawnictwo Politechniki Wrocławskiej, Wrocław
- Zipser T, Mlek M (2005) Modelowe próby interpretacji prawa Zipfa w systemie osadniczym, *Studia KPZK*, vol CXIV. PAN, Warszawa