Project 2: Species' Importance in a Food Web

Group 4: Dongjin Li, JunSeok Yeo, Evelyn Zhang, Yi Zheng.

MATH 381 A Win16. March 1st 2016

Contents

I. Introduction

Our planet is experiencing the 6th mass extinction of animals [1] and therefore a pressing ecological challenge is to predict the effects of each species' extinction [2]. Given a predator/prey food web, ecologists want to figure out the importance of each species' role in their ecosystem. Species are not isolated, but connected with their preys and predators in a sophisticated network. Thus it is impossible to find out which species is crucial just by observation and enumeration.

Our group works on simulating and analyzing the food web structure. First of all, by converting the whole predator-prey structure into a directed graph, we were able to rank the species by its degree. Then with the application of the Markov Chains model, we are able to find out how dominant each species is - in other words, its importance - in this food web. Our third method, which is also the method we heavily work on, is inspired by Google™ Search's algorithm PageRank™ [3], which is a way of measuring the importance of website pages based on the number and quality of links to a page. PageRank™ is basically a modification and optimization of the Markov Chains model. Back to our problem, intuitively, species with more interactions in the network should be considered more important. Therefore, in our model, we measure the importance of each species in its ecosystem based on the number and importance of other species related to it.

II. Background

An early idea for this project is to optimize the web organization of MyUW page. We found that some links on MyUW page are rarely used and some are used a lot. We want to construct a Markov chain of these page links and rank them by their importance so that the links we use more frequently are displayed at the top of the page. However, it was difficult for us to access the database of MyUW. Hence, we tried to find some similar ranking problems such sports teams' ranking, blogs ranking, but the datasets do not fit our model well. Eventually, we came up with this food web problem which fits our basic mathematical model well and has accessible database.

The number of endangered animals has been increasing rapidly in recent decades, predicting the effects of species' extinctions is essential. However, it is uneconomical and implausible to protect all the species in the ecosystems since even one ecosystem contains uncountably many species. For this reason, the measure of relative importance of species is an important topic. Because species are mutually dependent and are connected in a complex network of relationship known as food web, the loss of a single species can cause multiple co-extinctions, that is, if a species is on the edge of extinction, chain reactions occur and its predators, as well as preys are also threatened with extinction. What we want to do is to order species according to their importance for co-extinctions, in other words, finding the most important species whose extinction will result in fastest collapse of the ecological network in a certain area.

There are numerous measurements what species are more important and should be protected more than other species. There are several methods that biologists or ecologists typically use for this measurement. For example, one approach is to measure the number of connections between species [6]. In this approach, the species that has most connections with other species are considered as pivotal species in their ecosystem and should be protected first. Moreover, other approach is to measure how much the species are located in the center of ecosystem[7]. In other words, the more central a species is located, the more important the specie is. Furthermore, the methods could be more sophisticated and complicated to measure the importance of species in ecosystems. For instance, one method uses PageRank™ to measure the importance of species in the ecosystem. PageRank™ is the algorithm originally used by Google™ to rank web pages according to their relative importance, and it lies on the heart of Google search engine [8]. This algorithm evaluates the importance of web pages in a recursive definition using Markov Chain and lists them as a ranking. In the similar respect, researchers use PageRank™ to measure the importance of species in ecosystem[9]. There are all kinds of various methods of measuring the importance of species in ecosystem other than these three examples, and the list of the various methods would be very long. However, the point is that these methods have merits and demerits. Thus, we decided to compare these various methods to figure out which method is the most suitable to measure species importance value.

III. The Model

Part I: Data and Format

In order to test the functionality of various measurements of species importance, appropriate ecosystem and its food web are required. From several online articles and databases, we found ecosystems and food webs of Coachella Valley[12], El Verde rainforest[13], Papua New Guinea[14], and Serengeti National Park[15]. Since the purpose of this project is to test various measurements in order to figure out the most appropriate method for ranking species' importance value, we want the ecosystem and its food web be large enough with various species including plant, insects, mammal, reptile, and birds. Furthermore, it would be better if the ecosystem contains from small organisms like flies and rat to large organisms like elephant and giraffe. Coachella Valley is a desert, so there is not so much variety of animal. On the other hand, El Verde rainforest and Papua New Guinea have various kinds of animals because they are rainforests with numerous resources for all kind of animals. However, most animals in this ecosystem are small because small sized animals are usually proper to adapt to the environment of rainforest. For these reasons, we chose the ecosystem of Serengeti National Park as our database. Serengeti has renowned various kinds of animals because it has highland montane forest and treeless plains with abundant small plants. It has very small animal like flies to really huge animal like elephant. Furthermore, all the animals and plants are closely related to one another because mammals there are experiencing the largest terrestrial mammal migration in the world. This is why Serengeti national park is called as one of the seven natural wonders of Africa and one of the ten natural travel wonders of the world. Therefore, the ecosystem of Serengeti National Park[15] is chosen to be the date to test functionality of various measurements, and its food web is converted to data as an input for code.

After the conversion, we have two txt files for the database:

'Label.txt' contains all kinds of species and all species are number coded. The first element is the name of the species and the second element is its corresponding number label. These two elements are divided by a semicolon. The format is as following, for the full file, refer to appendix A:

Decaying material;1

Plant juices;2

Fruits and nectar;3

Grains, seeds;4

Grass and herbs;5

…

Nile crocodile;91

Maasai giraffe;92

Black rhinoceros;93

Hippopotamus;94

African elephant;95

'Predator-Prey Relations.txt' is the file that explains consumer-resource (i.e. predator-prey) links between all kinds of species in our dataset. The first number is the label for the predator and the second number is the label for the prey. These two elements are also divided by a semicolon. The format is as following, for the full file, refer to appendix A:

2;1

3;1

4;1

5;1

6;1

7;1

8;1

8;2

9;1

…

92;7

93;5

93;7

94;5

95;5

Part II: The Mathematical Model

First, we created a graph in which each vertex represents a species. Species receive nutrients and energy from primary producers and provide food resources for its predators [10][11]. We link the vertices based on their predator-prey relationship. More specifically, nutrients and energy moves through the resource-consumer network in ecosystem (i.e. the nutrients and energy moves from resources (preys) to consumers (predators)). So we can say that a consumer could survive based on the existence of its resources (i.e. resource supports its consumer). Therefore we constructed a directed graph, with edges directed from each consumer to its corresponding resources. Here is an example of a small food web graph with only five animals but in our actual dataset, we have 95 species in total.

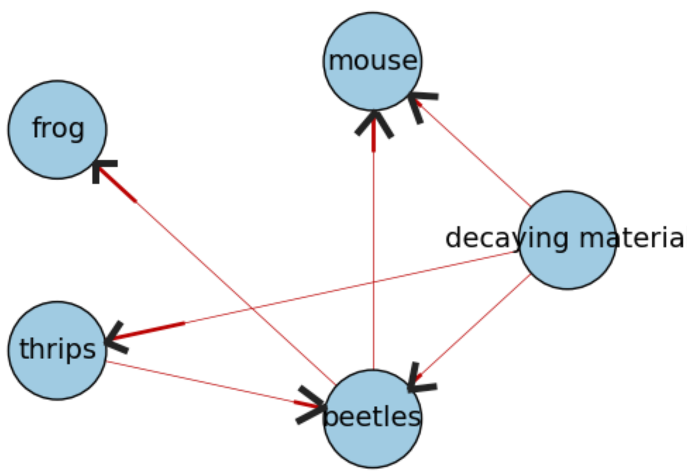


Figure 1. A Small Example of Graph with 5 Vertices

Method 1. Counting Degree

After constructing the graph, we counted the degree of each vertex and ranked the species' importance by comparing their degrees, in other words, how many connections they have. In the example, decaying material has a degree of 3, thrips has degree 1, beetle has degree 2, mouse has degree 0 and frog has degree 0 as well. A higher larger degree indicates the species is more important.

Based on the example, we ranked the species in decreasing order of degree as following:

decaying material, beetles, thrips, mouse, frog

In the same manner, we constructed the graph with whole 95 species and made a rank with respect to the degree of each vertex. Here is a shorten version of the result for this method. Please check Appendix B for the full result.

Top Five	Rank	Bottom Five	Rank
Crickets	1	African elephant	91
Fat or tree mouse	2	Domestic cat	92
Grains	3	Marabou stork	93
Decaying material	4	Domestic dog	94
Rove and ground beetles	5	Hippopotamus	95

Table 1: Partial Result from the Counting Degree Method.

Based on the result, we can say that the top five species on the list make the most connections with other species and play essential roles in this ecosystem. Take crickets as an example, they eat plant materials, fungi and some other insects and worms. On the other hand, they serve as food for many species of birds, reptiles, spiders, wasps, and rodents [18]. Crickets are the predator and prey of various species and therefore they have a huge number of connections. Hence it is at the very top of the list. Take hippopotamus as another example, it basically feeds on grass and fruits, plus only Nile crocodiles, lions and spotted hyenas are known to prey on young hippos[19]. Hippos have very few connections with other species and therefore it is at the bottom of the list.

Method 2: Markov Chains

We constructed a Markov chain with transition matrix S and defined the states of the chain as all kinds of species in our dataset. A species is important if it supports (directly or indirectly) other species that are in turn important (i.e. species are important if they are resources of other important species)[9] . In order to compare species' importance value, we want to find out the probability the Markov Chain stays in the state. Here we take use of Limiting Distribution, which describes the long-term behavior of Markov Chains, in particular, it shows the probability of the Markov Chain stays in each state. In our case, the limiting distribution represents the possibility of a randomly chosen species finally get to another species that provides nutrients to the original chosen species(directly or indirectly). Higher probability of a certain species, say A, means that the species we randomly choose are more likely to receive energy and nutrients (directly or indirectly) from that species A. Thus, if the species with high probability goes to extinction, then there is a high possibility that numbers of other species will lose their energy and nutrients resources and eventually die, resulting in a collapse of the ecological network in a that area. Therefore, the species with high probability can be considered as of high importance.

To solve for the Limiting Distribution, we construct a matrix A in which each entry a_{ij} is an indicator telling if the species i eats species j, that is, if species i receives energy and nutrients directly from species j.

$$a_{ij} = \begin{cases} 1, & if\ \ i\ eats\ j \\ 0, & otherwise \end{cases}$$

Then we build a transition matrix T in which each element t_{ij} represents the fraction of possibility assigned to a connection between two species. We assume that if a predator has several kinds of preys, it is equally possible for the predator to pick each of its preys as food. Thus, t_{ij} is given by $t_{ij} = a_{ij}/\sum_j a_{ij}$. As long as the matrix S is ergodic, we can solve the problem by finding the normalized left eigenvector corresponding to the eigenvalue of 1 of the matrix T.

We used python to generate the Limiting Distribution and all the commented code can be found in appendix C. For the database of 95 species, using the original graph, which has edges directed from each consumer to its corresponding resources, we obtained a result like this:

Top Five	Rank
African marsh owl	1
Nile crocodile	2
Black backed jackal	3
Rove and ground beetles	4
Dwarf mongoose	5

Table2: Partial Result from Markov Chain Method.

We can see that high ranking animals are the ones that consumes most other species. However, extinction of preys directly lead to dramatic decrease of predator population and consequently raises the risk of its predators' extinction. Thus species that are preys to various other species are considered more important in this case. Therefore, after thorough consideration, we decided to reverse the direction of our directed graph in order to put species that are popular prey to top of the list.

In our small example, the original food web did not have any closed circle and thus we were not able to convert it into a Markov Chain problem. Therefore we added connections between each species to the decaying material, which decomposes pretty much all other species and is food resource to plants, insects and some other species. Adding these connections makes the transition matrix ergodic, thus allowing us to construct Markov Chains and calculate Limiting Distributions. The modification on the connections is shown in figure 2. In figure 3, we reversed the directions of the connections to give popular preys, instead of consumers, higher rankings, our reasoning has been explained on page 10.

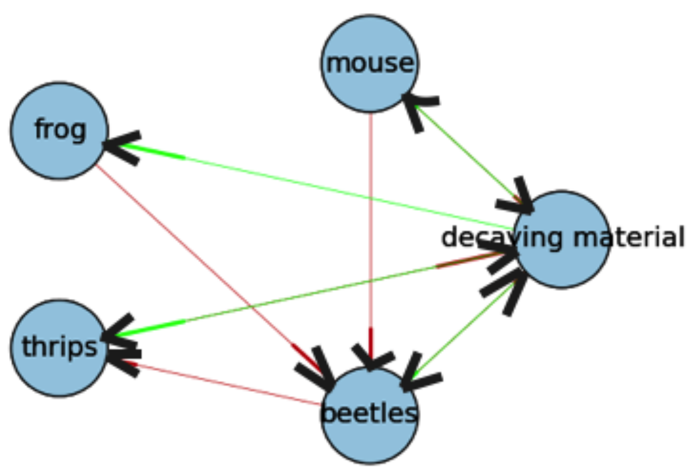


Figure 2. Add Edge from Each Eode to Decaying Material (in green)

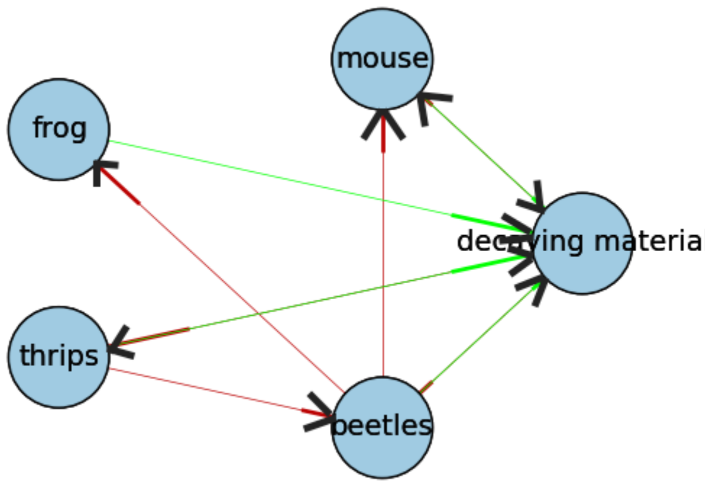


Figure 3. Reverse the direction of all edges

In the example, we defined the states of the chain as {decaying material, thrips, beetles, mouse, frog}. Then the transition-probability matrix looks like this:

T =

	decaying material	**thrips**	**beetles**	**mouse**	**frog**
decaying material	0	1/4	1/4	1/4	1/4
thrips	1	0	0	0	0
beetles	1/2	1/2	0	0	0
Mouse	1/2	0	1/2	0	0
frog	0	0	1	0	0

Since this Markov chain is ergodic, we could compute the limiting distribution by getting the normalized eigenvector corresponding to the eigenvalue 1.

Here is the Limiting Distribution we got:

{0.3721, 0.2093, 0.2326, 0.0930, 0.0930}

So, based on the definition of limiting distribution, the ranking is as following:

decaying material, beetles, thrips, mouse, frog

In the same manner, we constructed a (95, 95) transition matrix with each stage representing our 95 species. Here is a shorten version of the result and please refer to appendix B for the full result.

Top Five	Rank	Limiting Distribution	Bottom Five	Rank	Limiting Distribution
Decaying material	1	(0.2752544458j)	African marsh owl	91	(0.002928238j)
Grass and herbs	2	(0.0487459105j)	Black backed jackal	92	(0.002928238j)
Fruits and nectar	3	(0.0427373528j)	Domestic cat	93	(0.002928238j)
Trees and shrubs	4	(0.0387520153j)	Marabou stork	94	(0.002928238j)
Crickets	5	(0.0309442659j)	Domestic dog	95	(0.002928238j)

Table 3: Partial Result from the Markov Chain Method.

From this result we can see that the top 5 species are food resources to a lot of other species. Generally, the last five species are all omnivores which have wide ranges of food sources. Furthermore, looking closely at the database, we found out that they do not any connection with other species. Thus they have small limiting distribution values and low ranks. Notice that these low ranking states all have equally low Limiting Distribution value, which means that the extinction of any of these species would not have dramatic effect on other species. The differences of rankings are negligible when the rankings are low, in our case, any species with ranking greater than 77 have the same limiting distribution.

Method 3: PageRank

PageRank is an algorithm originally used to rank the websites by its relative importance in a given set of websites. The Wikipedia article on PageRank https://en.wikipedia.org/wiki/PageRank says "PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. More important websites are likely to receive links from other websites." The World Wide Web is a directed network where websites (nodes) are connected with each other by hyperlinks. PageRank algorithm is based on this web graph in which web pages are nodes and hyperlinks are edges. The basic idea of this algorithm is pretty similar to Markov Chain model. It assigns a numerical value to each web page in the hyperlinked set of websites to represents its importance. Mathematically, we can construct a matrix A in which each entry a_{ij} is an indicator telling whether j links to i (exiting page i to enter page j).

$$a_{ij} = \begin{cases} 1, & if\ \ j\ links\ to\ i \\ 0, & otherwise \end{cases}$$

Then to solve the recursive problem, by assuming that the probability of clicking we build a matrix S in which each element s_{ij} represents the fraction of importance assigned to a link and $s_{ij} = a_{ij} / \sum_j a_{ij}$. Similarly, when the matrix S is ergodic, we can solve the problem by finding the normalized left eigenvector corresponding to $\lambda = 1$ of the matrix S.

Here we use the PageRank algorithm to rank the importance of species for the food web in Serengeti National Park. Nutrients move from one species to another in a food web through feeding links. To survive, species must be able to get energy and matter from primary producers through some pathway in the network. As mentioned in method 2, we define a species as important if it supports (directly or indirectly) other species that are in turn important, which is similar to web page ranking.

The PageRank theory holds that an imaginary surfer who is randomly clicking on links will eventually stop clicking. So to be more practical, it adds a damping factor. In PageRank algorithm, the probability that the surfer will continue clicking on links the is the damping factor d. Then (1-d) is the probability that the surfer browsing the web can decide to move directly to another random page. In our case, the damping factor could actually mimic the probability that a certain kind of species dying before natural death or being killed (because of weather, disease or other random events), instead of being eaten by their predators as indicated in the food web, will be recycled into the food web and randomly become other species. For example, if a rabbit is killed by a deadly disease and its body is then eaten by a tiger, the rabbit would be considered to become the tiger. To be more specific, in this case, the conversion from a rabbit to a tiger is like a jump rather than a normal process happening in food web.

Various studies of PageRank have tested different damping factors, but generally it is assumed that the damping factor will be set around 0.85. We have tried different damping factors but the result does not change at all. So we stick with d=0.85. Applying this damping factor d, we modified our matrix in method 2 into a new Google matrix G with entries $G_{ij}=d*s_{ij}+(1-d)/N$, where N is the total number of species we picked (number of vertices in our graph).

For the same example as used in method 2, we constructed the new Google matrix G:

G =

	decaying material	**thrips**	**beetles**	**mouse**	**frog**
decaying material	0.03	0.2425	0.2425	0.2425	0.2425
thrips	0.88	0.03	0.03	0.03	0.03
beetles	0.455	0.455	0.03	0.03	0.03
Mouse	0.455	0.03	0.455	0.03	0.03
frog	0.03	0.03	0.88	0.03	0.03

Computing the limiting distribution, we could get:

{0.3494, 0.2050, 0.2371, 0.1042, 0.1042}

So, based on the limiting distribution, the ranking is as following:

decaying material, beetles, thrips, mouse, frog

The decaying material, again, is the most important species because it is connected to all the other four species.

Top Five	Rank	Limiting Distribution	Bottom Five	Rank	Limiting Distribution
Decaying material	1	(0.229585735j)	African elephant	91	(0.003654988598j)
Grass and herbs	2	(0.046954991j)	Domestic cat	92	(0.003654988598j)
Fruits and nectar	3	(0.040583501j)	Marabou stork	93	(0.003654988598j)
Trees and shrubs	4	(0.037594827j)	Domestic dog	94	(0.003654988598j)
Grains, seeds	5	(0.029689179j)	Hippopotamus	95	(0.003654988598j)

Table 4: Partial Result from the PageRank Method.

We can see that the ranking result from the PageRank method did not change much from the Markov Chains method. There are some minor differences of ranking but no large variations. For both results, when looking at the limiting distribution values, we see that there is a huge gap between the species ranked first and the one ranked second. Note that the species at the top of the list is decaying material, which decomposes all other species and thus is connected with all other species. Therefore there the probability of the Markov Chain staying at this state is considerably higher than other species.

Method 4: Closeness Centrality

In connected graphs there is a natural distance metric between all pairs of nodes, defined by the length of their shortest paths. The farness of a node x is defined as the sum of its distances from all other nodes, and its closeness was defined by Bavelas as the reciprocal of the farness. [16][17]

So the the closeness centrality of a node u is the reciprocal of the sum of the shortest path distances from u to all n-1 other nodes. Since the sum of the shortest path depends on the number of nodes in graph, closeness is normalized by the sum of minimum possible distance n-1.

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)},$$

where d(v, u) is the shortest-path distance between v and u, and n is the number of nodes in the graph.

The higher values of closeness indicate higher centrality

In our example, we have

d(frog, beetles) = 1, d(frog, thrips) = 2, d(frog, mouse) = 3, d(frog, decaying material) = 2

Hence, C(frog) = 4 / (1+2+3+2) = 0.5

In this way, we could also get

C(decaying material) = 4/4 = 1

C(thrips) = 4/7 = 0.57

C(beetles) = 4/6 = 0.67

C(mouse) = 4/6 = 0.67

Therefore, we could get the rank by closeness:

decaying material, beetles, mouse, thrips, frog

Similarly, we wrote a python program to compute the closeness of vertices in the graph and ranked the species. Here is a shorten graph of out result.

(Code can be found in appendix C and full list of result can be found in appendix B.)

Top Five	Rank	Closeness	Bottom Five	Rank	Closeness
Decaying material	1	1	Yellow-throated Sandgrouse	91	0.3369175
Black backed jackal	2	0.57668711	Dwarf Epauletted Bat	92	0.3369175
Fat or tree mouse	3	0.52808988	Buffalo	93	0.3369175
Creek rat	4	0.52513966	Bush hyrax	94	0.3369175
Katydids	5	0.51933701	Hippopotamus	95	0.3369175

Table 5: Partial Result from the Closeness Centrality Method.

As expected, decaying material, which is connected with all species, is once again on the top of the list. Then recursively, species closely related to high ranking species receive high rankings. One drawback of this method is that since it is measured based on vertices' distance to the center of the graph, several species have the same closeness and the difference between these species are filtered out.

IV. Conclusion

All four measurements using degree of vertexes, markov chain, Pagerank, and centrality in the graph shows general tendency for results. Top ranks are commonly occupied by primary producers or small animals. Primary producers, plants, produces energy or resources from the environments by using photosynthesis, decaying materials, or nutrients in soils or primary consumers. Small animals eat plants directly and are eaten by higher trophic level of consumers. Therefore, the primary producers and consumers are usually ranked highly in all of our measurements because they usually provide most nutrients and energy to other animals in the ecosystem. On the other hand, the large animals that occupies the top trophic level in the food web tends to be located in the bottom of ranking since they are usually not eaten by other predators in the food web, and they do not provide so much nutrients or energy to other animals in the ecosystem. Furthermore, it could be noted that large herbivore or domestic animals occupy the very bottom of the ranking. Domestic animals has very low interaction with wild animals since they are raised and feeded by human beings. Large herbivore would also have influence on the ecosystem because they would only eat plants and are not likely be eaten by other predators. Thus, these general tendencies of the results make sense. If primary producers or consumers go to extinction, consumers of the higher trophic level that prey on these primary producers or consumers mainly will starve to death sequentially because their foods are gone, and it could end up the mass extinction in the ecosystem. In contrary, the extinction of high trophic level consumers like elephant or hippopotamus, would not lead to several results as if the extinction of primary producers or consumers because few or no animal prey on them. Therefore, all the measurements agree with general tendency of the important species in the ecosystem.

However, there are still differences in four measurements since they evaluate the importance of species in different perspectives. First degree of vertexes method is simple but useful method to measure the importance of species. However, it only measures direct relations between species, so it would not be able to account for the secondary effects that species influence on other species that are not directly connected sequentially by affecting bridging species. Second Markov chain and third Pagerank methods are very similar to each other because both methods use Markov chain to explains secondary effect, and the only major difference between two methods is that Pagerank method has damping factor while Markov chain method does not. Only defect of these methods is that these methods would give more importance on preys or predators. In other words, it is hard for this methods to give equal importance to preys and predators since these methods use a direct graph. These two methods successfully includes secondary effects by finding limiting distribution of Markov chain. The last centrality method implies secondary effect differently from second and third methods by evaluating how much a species is located on the center of food web, and this method equally give importance to prey and predators. However, this method would be less effective for evaluating secondary effect than the second and their methods because the second and third methods postulate the secondary effect iteratively by using Markov chain while the fourth methods just measure how one species is closely connected to all other species in the food web.

Therefore, all four methods have merits and demerits, and the appropriate method would be dependent on the circumstance of food web. If the direct relation is most significant, first method would be better. On the other hand, in the case that the secondary effect has significant impact, the second or third methods would be better. Last, the fourth method would be better if consideration of secondary effect is needed with equal importance to preys and predators.

V. Bibliorgraphy

[1] Kaplan, Sarah. *"Earth is on brink of a sixth mass extinction, scientists say, and it's humans' fault"* Washington Post. June 22, 2015.

[2] Claude Gascon, Thomas M. Books. *"The Importance Benefits of Species."* Current Biology. Volume 25, Issue 10. May 18, 2015.

[3] Sullivan, Danny. *"What is Google PageRank? A guide for Searchers & Webmasters."* Search Engine Land. April 26, 2007.

[4] Zavaleta E, Hulvey K *"Realistic species losses disproportionately reduce grassland resistance to biological invaders."* Science 306: 1175–1177. 2004.

[5] Montoya J, Pimm S, Sole R *"Ecological networks and their fragility."* Nature 442: 259–264. 2006.

[6] Dunne J, Williams R, Martinez N. *"Network structure and biodiversity loss in food webs: Robustness increases with connectance."* Ecol Lett 5: 558–567. 2002.

[7] Jordan F, Benedek Z, Podani J. *"Quantifying positional importance in food webs: A comparison of centrality indices."* Ecol Model 205: 270–275. 2007.

[9] Allesina S, Pascual M *"Googling Food Webs: Can an Eigenvector Measure Species' Importance for Coextinctions?"* PLoS Comput Biol 5(9): e1000494. doi:10.1371/journal.pcbi.1000494. 2009

[10] Allesina S, Bodini A.*"Who dominates whom in the ecosystem? Energy flow bottlenecks and cascading extinctions."* J Theor Biol 230: 351–358. 2004

[11] Allesina S, Bodini A, Pascual M *"Functional links and robustness in food webs."* Phil Trans Roy Soc B 364: 1701–1709. 2009

[12] Polis G. Complex trophic interactions in deserts: an empirical critique of food-web theory. 1991

[13] Reagan D, Waide R (1996) The food web of a tropical rain forest University of Chicago Press.

[14] Novotny V, Miller SE, Baje L, Balagawi S, Basset Y, et al. (2010) Guild-specific patterns of species richness and host specialization in plant–herbivore food webs from a tropical forest. J Anim Ecol 79

[15] de Visser, S. N., Freymann, B. P. and Olff, H. (2011), The Serengeti food web: empirical quantification and analysis of topological changes under increasing human impact. Journal of Animal Ecology, 80: 484–494. doi:10.1111/j.1365-2656.2010.01787.

[16] Alex Bavelas. Communication patterns in task-oriented groups. J. Acoust. Soc. Am 1950.

[17]Sabidussi, G. "The centrality index of a graph". Psychometrika 31: 581–603. 1966.

[18] “What do Crickets Eat?” Source: http://www.whatdocricketseat.info/ 1999.

[19]Kingdon, J. “*East African Mammals: An Atlas of Evolution in Africa Volume 3, Part B: Large Mammals*.” University Of Chicago Press. pp. 256–77. ISBN 0-226-43722-1. 1988.

VI. Appendix

1. Appendix A: Database

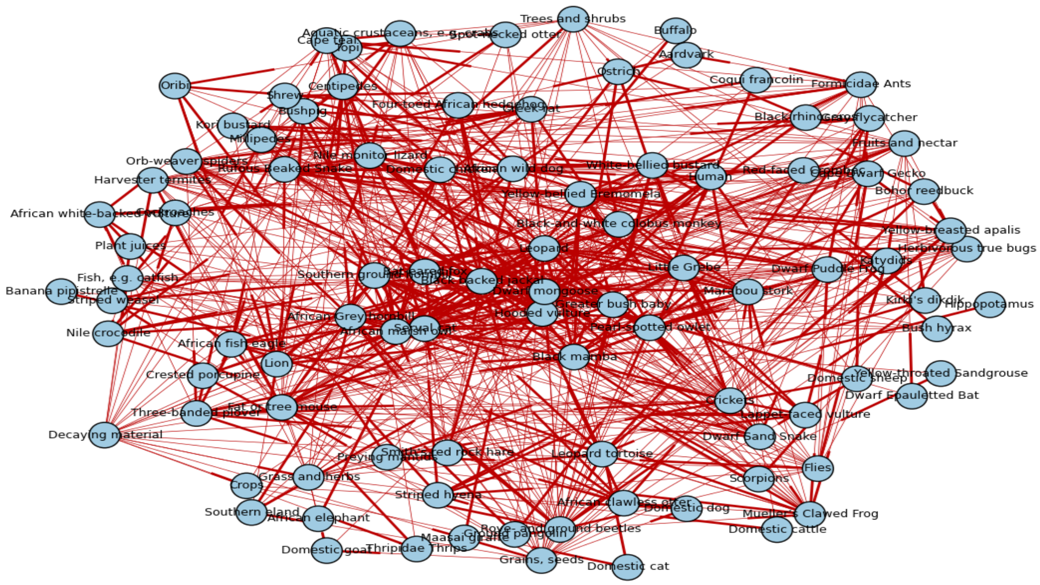
   (a) Labels of each variable - “Label.txt”:

Decaying material	1	Cape teal	33	Kori bustard	65
Plant juices	2	Cape Dwarf Gecko	34	Black-and-white colobus monkey	66
Fruits and nectar	3	Aquatic crustaceans, e.g. crabs	35	Black backed jackal	67
Grains, seeds	4	Creek rat	36	Ground pangolin	68
Grass and herbs	5	Fish, e.g. catfish	37	Serval cat	69
Crops	6	Dwarf Sand Snake	38	African clawless otter	70
Trees and shrubs	7	Pearl-spotted owlet	39	Kirki's dikdik	71
Thripidae Thrips	8	Yellow-throated Sandgrouse	40	Crested porcupine	72
Rove and ground beetles	9	Striped weasel	41	Oribi	73
Orb-weaver spiders	10	African Grey hornbill	42	African wild dog	74
Formicidae Ants	11	Coqui francolin	43	Black mamba	75
Harvester termites	12	Rufous Beaked Snake	44	Domestic dog	76
Flies	13	Four-toed African hedgehog	45	Striped hyena	77
Crickets	14	Little Grebe	46	Domestic goat	78
Herbivorous true bugs	15	White-bellied bustard	47	Aardvark	79
Katydids	16	Bush hyrax	48	Leopard	80
Preying mantids	17	African marsh owl	49	Domestic sheep	81
Centipedes	18	Dwarf mongoose	50	Bohor reedbuck	82
Cockroaches	19	Hooded vulture	51	Bushpig	83
Dwarf Puddle Frog	20	Smith's red rock hare	52	Human	84
Millipedes	21	African fish eagle	53	Ostrich	85
Shrew	22	Domestic chicken	54	Lion	86
Banana pipistrelle	23	Greater bush baby	55	Topi	87
Scorpions	24	Marabou stork	56	Buffalo	88
Gray flycatcher	25	Southern ground hornbill	57	Southern eland	89
Yellow-breasted apalis	26	Domestic cat	58	Domestic cattle	90
Yellow-bellied Eremomela	27	Bat-eared fox	59	Nile crocodile	91
Red-faced Crombec	28	Spot-necked otter	60	Maasai giraffe	92
Three-banded plover	29	Leopard tortoise	61	Black rhinoceros	93
Dwarf Epauletted Bat	30	African white-backed vulture	62	Hippopotamus	94
Fat or tree mouse	31	Lappet-faced vulture	63	African elephant	95
Mueller's Clawed Frog	32	Nile monitor lizard	64		

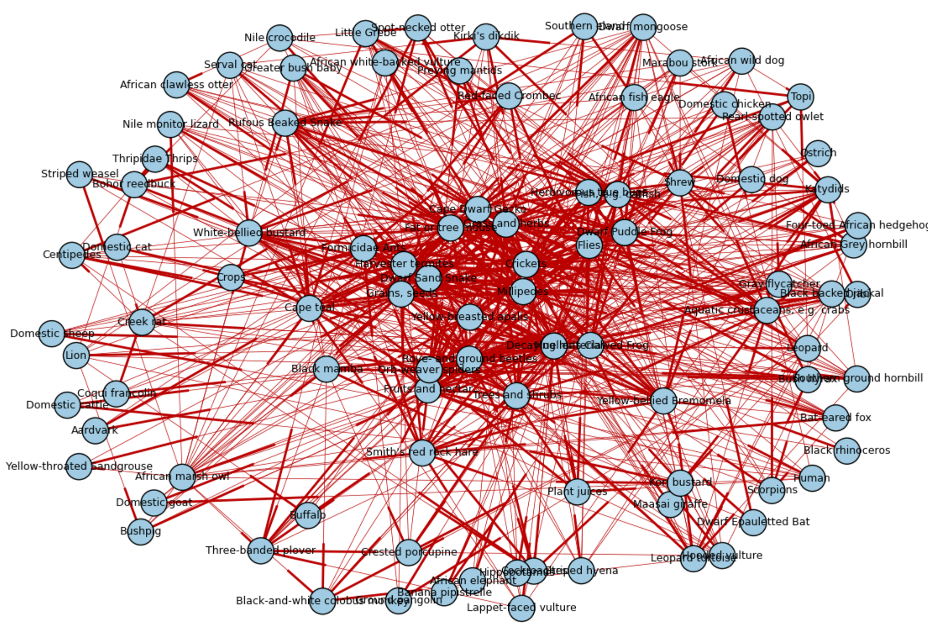
(b) "Predator-Prey Relations.txt"

2;1	18;31	28;3	37;1	44;27	49;52	53;34	59;21	67;1	73;3
3;1	19;1	28;4	37;37	44;28	49;55	53;37	59;24	67;9	73;4
4;1	20;9	29;9	38;20	44;31	49;59	53;38	59;27	67;10	73;5
5;1	20;10	29;10	38;27	44;33	49;60	53;42	59;31	67;11	73;7
6;1	20;11	29;13	38;31	44;34	49;64	53;44	59;34	67;12	74;26
7;1	20;12	29;14	38;37	44;38	49;65	53;61	59;38	67;13	74;27
8;1	20;13	29;15	38;38	44;44	49;66	54;4	59;44	67;14	74;31
8;2	20;14	29;21	38;44	45;4	49;75	54;5	59;47	67;19	74;47
9;1	21;1	29;35	39;9	45;9	50;3	54;6	59;52	67;20	74;52
9;8	21;3	30;3	39;13	45;10	50;9	54;14	60;35	67;21	74;71
9;9	22;9	31;1	39;14	45;14	50;10	55;2	60;37	67;22	74;72
9;12	22;10	31;3	39;15	45;22	50;11	55;3	61;4	67;24	74;73
9;13	22;11	31;4	39;16	45;31	50;12	55;4	61;5	67;31	74;79
9;14	22;14	31;5	39;20	46;9	50;13	55;9	61;9	67;32	74;82
9;15	22;15	31;7	39;21	46;13	50;14	55;14	61;14	67;34	74;87
9;19	22;21	31;9	39;22	46;20	50;16	55;21	61;20	67;35	75;20
10;9	23;13	31;11	39;23	46;21	50;18	55;27	61;21	67;37	75;26
10;10	24;10	31;12	39;24	46;22	50;20	55;31	61;32	67;38	75;27
10;11	24;13	31;14	39;25	46;31	50;21	55;33	61;34	67;44	75;28
10;13	24;15	31;19	39;26	46;32	50;22	55;34	61;35	67;45	75;29
10;15	24;24	31;31	39;27	46;34	50;30	55;52	61;37	67;46	75;31
10;17	25;9	32;1	39;31	46;35	50;31	56;12	62;27	67;47	75;32
11;1	25;11	32;9	39;32	46;37	50;32	56;14	62;28	67;52	75;33
11;7	25;12	32;10	39;33	46;38	50;34	56;16	62;29	67;61	75;34
11;12	25;13	32;14	39;34	47;3	50;35	56;20	62;31	67;65	75;37
12;1	25;14	32;16	39;38	47;4	50;36	56;22	62;33	68;11	75;38
12;5	25;21	32;20	39;44	47;9	50;37	56;31	62;34	68;12	75;44
13;1	26;3	32;32	40;4	47;10	50;38	56;32	62;37	69;31	75;48
13;3	26;4	32;37	41;26	47;14	50;44	56;33	62;52	69;33	75;60
14;1	26;9	33;3	41;27	47;16	50;47	56;37	63;27	69;34	75;66
14;2	26;10	33;4	41;31	47;20	50;48	56;38	63;29	69;37	75;72
14;5	26;11	33;7	41;33	47;21	50;52	56;44	63;31	69;38	76;4
14;7	26;12	33;11	42;3	47;31	51;22	57;9	63;33	69;43	76;6
14;14	26;13	33;14	42;4	47;34	51;27	57;10	63;34	69;47	77;9
15;2	26;14	33;21	42;5	47;38	51;28	57;13	63;37	69;48	77;12
16;1	26;15	34;10	42;9	47;44	51;29	57;14	63;52	69;52	77;13
16;3	26;16	34;11	42;10	48;7	51;31	57;16	64;9	69;71	77;14
16;4	26;34	34;13	42;13	49;22	51;33	57;20	64;14	70;20	77;15
16;13	27;3	34;14	42;14	49;25	51;47	57;21	64;16	70;21	77;16
16;14	27;4	34;15	42;15	49;26	51;48	57;22	64;19	70;31	77;35
16;15	27;9	35;1	42;16	49;27	51;52	57;31	64;20	70;32	78;4
16;21	27;10	35;21	42;20	49;28	51;60	57;32	64;31	70;35	78;5
17;10	27;11	36;1	42;33	49;29	51;65	57;34	64;32	70;37	79;11
17;13	27;12	36;9	42;35	49;31	51;66	57;61	64;33	70;42	79;12
17;14	27;13	36;10	42;37	49;33	51;75	57;75	64;35	70;61	80;20
17;15	27;14	36;11	43;4	49;34	52;5	58;6	64;38	71;3	80;31
18;8	27;15	36;12	43;12	49;40	52;7	58;36	65;3	71;4	80;32
18;14	27;16	36;14	44;11	49;41	53;20	59;11	65;34	71;5	80;34
18;18	27;17	36;18	44;12	49;45	53;22	59;12	66;1	71;7	80;37
18;19	27;20	36;19	44;20	49;46	53;29	59;14	66;3	72;3	80;47
18;21	27;26	36;21	44;21	49;47	53;31	59;18	66;7	72;4	80;48
18;22	27;34		44;26	49;50	53;32		66;21		80;52

(c) Predator-prey Directed Gragh



Predator-prey Directed Graph – reversed.



2. Appendix B: Result

(a) Ranking by **Degree**:

Crickets	31	Plant juices	5	Leopard	1
Fat or tree mouse	27	Bush hyrax	5	Dwarf Epauletted Bat	1
Grains, seeds	25	Leopard tortoise	5	Domestic goat	1
Decaying material	23	Spot-necked otter	4	Bat-eared fox	1
Rove- and ground beetles	23	Scorpions	4	Buffalo	1
Fruits and nectar	22	Crops	4	Nile monitor lizard	1
Grass and herbs	21	Topi	4	Striped weasel	1
Cape Dwarf Gecko	20	Centipedes	4	Aardvark	1
Millipedes	20	Kirki's dikdik	4	Domestic sheep	1
Flies	19	Black-and-white colobus monkey	4	Dwarf mongoose	1
Harvester termites	18	Bohor reedbuck	3	Coqui francolin	1
Dwarf Puddle Frog	18	Black mamba	3	African wild dog	0
Fish, e.g. catfish	18	Greater bush baby	3	Maasai giraffe	0
Formicidae Ants	17	Oribi	3	Lappet-faced vulture	0
Orb-weaver spiders	17	Crested porcupine	3	African fish eagle	0
Trees and shrubs	16	Preying mantids	2	Black rhinoceros	0
Cape teal	14	Gray flycatcher	2	Hooded vulture	0
Dwarf Sand Snake	13	Southern eland	2	Serval cat	0
Herbivorous true bugs	13	Southern ground hornbill	2	African white-backed vulture	0
Mueller's Clawed Frog	13	African Grey hornbill	2	Ostrich	0
Yellow-bellied Eremomela	12	Thripidae Thrips	2	Lion	0
Katydids	11	Creek rat	2	Human	0
Smith's red rock hare	11	Little Grebe	2	Pearl-spotted owlet	0
Shrew	11	Four-toed African hedgehog	2	Striped hyena	0
Rufous Beaked Snake	10	Ground pangolin	1	African marsh owl	0
Aquatic crustaceans, e.g. crabs	10	Nile crocodile	1	Black backed jackal	0
White-bellied bustard	8	Domestic cattle	1	African elephant	0
Yellow-breasted apalis	7	Domestic chicken	1	Domestic cat	0
Cockroaches	6	Banana pipistrelle	1	Marabou stork	0
Three-banded plover	6	Bushpig	1	Domestic dog	0

Kori bustard	5	Yellow-throated Sandgrouse	1	Hippopotamus	0
Red-faced Crombec	5	African clawless otter	1		

(b) Ranking by **Markov Chain** (limiting distribution):

Decaying material	(0.27525444588-0j)	African Grey hornbill	(0.00361276213853-0j)
Grass and herbs	(0.0487459105396-0j)	Crested porcupine	(0.00359211624407-0j)
Fruits and nectar	(0.042737352893-0j)	Black mamba	(0.00354462829278-0j)
Trees and shrubs	(0.0387520153047-0j)	Southern eland	(0.00342973331489-0j)
Crickets	(0.030944265904-0j)	Southern ground hornbill	(0.00342973331489-0j)
Plant juices	(0.0306372528376-0j)	Nile crocodile	(0.00334655861253-0j)
Grains, seeds	(0.0285222267892-0j)	African clawless otter	(0.00334655861253-0j)
Harvester termites	(0.0244014901917-0j)	Leopard	(0.00334655861253-0j)
Flies	(0.02292048512-0j)	Domestic cattle	(0.00325359865107-0j)
Fish, e.g. catfish	(0.022066363786-0j)	Domestic chicken	(0.00325359865107-0j)
Formicidae Ants	(0.0203572766211-0j)	Bushpig	(0.00325359865107-0j)
Rove- and ground beetles	(0.0194746758276-0j)	Domestic goat	(0.00325359865107-0j)
Orb-weaver spiders	(0.0182470317011-0j)	Buffalo	(0.00325359865107-0j)
Herbivorous true bugs	(0.0178656302995-0j)	Domestic sheep	(0.00325359865107-0j)
Millipedes	(0.0166899174447-0j)	Coqui francolin	(0.00322106266456-0j)
Fat or tree mouse	(0.013757137686-0j)	Gray flycatcher	(0.00320967134663-0j)
Cape Dwarf Gecko	(0.0106552816833-0j)	Aardvark	(0.00319444231196-0j)
Dwarf Puddle Frog	(0.00977007798035-0j)	Little Grebe	(0.00317268306723-0j)
Cockroaches	(0.0081702232555-0j)	Four-toed African hedgehog	(0.00317268306723-0j)
Aquatic crustaceans, e.g. crabs	(0.00790893824746-0j)	Ground pangolin	(0.00310437344978-0j)
Cape teal	(0.00786730304742-0j)	Banana pipistrelle	(0.0030823566168-0j)
Yellow-bellied Eremomela	(0.00779830084339-0j)	Dwarf Epauletted Bat	(0.00305555351579-0j)
Dwarf Sand Snake	(0.00733775374867-0j)	Yellow-throated Sandgrouse	(0.00305555351579-0j)
Katydids	(0.00702820411218-0j)	Bat-eared fox	(0.00305555351579-0j)
Crops	(0.0069952370998-0j)	Nile monitor lizard	(0.00305555351579-0j)
Mueller's Clawed Frog	(0.00679999272303-0j)	Striped weasel	(0.00305555351579-0j)
Preying mantids	(0.00652643222496-0j)	Dwarf mongoose	(0.00305555351579-0j)
Rufous Beaked Snake	(0.00645094878511-0j)	Black rhinoceros	(0.00292823878596-0j)

Thripidae Thrips	(0.00599470359319-0j)	African elephant	(0.00292823878596-0j)
Shrew	(0.00592493403024-0j)	Hippopotamus	(0.00292823878596-0j)
Smith's red rock hare	(0.00561162016544-0j)	African wild dog	(0.00292823878596-0j)
Yellow-breasted apalis	(0.00555590307411-0j)	Maasai giraffe	(0.00292823878596-0j)
Scorpions	(0.00457937166184-0j)	Lappet-faced vulture	(0.00292823878596-0j)
Three-banded plover	(0.00455289512228-0j)	African fish eagle	(0.00292823878596-0j)
Creek rat	(0.00451967290877-0j)	Hooded vulture	(0.00292823878596-0j)
White-bellied bustard	(0.00449545158098-0j)	Serval cat	(0.00292823878596-0j)
Centipedes	(0.00442491230146-0j)	African white-backed vulture	(0.00292823878596-0j)
Red-faced Crombec	(0.00440595083514-0j)	Ostrich	(0.00292823878596-0j)
Leopard tortoise	(0.00416985199339-0j)	Lion	(0.00292823878596-0j)
Topi	(0.00411425666745-0j)	Human	(0.00292823878596-0j)
Kirki's dikdik	(0.00398876071948-0j)	Pearl-spotted owlet	(0.00292823878596-0j)
Bush hyrax	(0.00397130046388-0j)	Striped hyena	(0.00292823878596-0j)
Bohor reedbuck	(0.00393812200363-0j)	African marsh owl	(0.00292823878596-0j)
Spot-necked otter	(0.00392066174803-0j)	Black backed jackal	(0.00292823878596-0j)
Kori bustard	(0.00389942673353-0j)	Domestic cat	(0.00292823878596-0j)
Oribi	(0.00369593684088-0j)	Marabou stork	(0.00292823878596-0j)
Black-and-white colobus monkey	(0.00367847658528-0j)	Domestic dog	(0.00292823878596-0j)
Greater bush baby	(0.00365000800617-0j)		

(c) Ranking by **PageRank** algorithm:

Species	Limiting Distribution	Name	Limiting Distribution
Decaying material	(0.22958573598-0j)	African Grey hornbill	(0.00437192866918-0j)
Grass and herbs	(0.0469549917779-0j)	Crested porcupine	(0.00434894408167-0j)
Fruits and nectar	(0.0405835019834-0j)	Black mamba	(0.0043025568414-0j)
Trees and shrubs	(0.0375948278044-0j)	Southern eland	(0.00418313317995-0j)
Grains, seeds	(0.0296891792582-0j)	Southern ground hornbill	(0.00418313317995-0j)
Crickets	(0.029420367112-0j)	Nile crocodile	(0.00408949773206-0j)
Plant juices	(0.0266233638399-0j)	African clawless otter	(0.00408949773206-0j)
Harvester termites	(0.023660472964-0j)	Leopard	(0.00408949773206-0j)
Flies	(0.022338250544-0j)	Domestic cattle	(0.00400018196562-0j)

Fish, e.g. catfish	(0.0206246533428-0j)	Domestic chicken	(0.00400018196562-0j)
Formicidae Ants	(0.0201946158718-0j)	Domestic goat	(0.00400018196562-0j)
Rove- and ground beetles	(0.0190536716455-0j)	Domestic sheep	(0.00400018196562-0j)
Orb-weaver spiders	(0.0178609095927-0j)	Bushpig	(0.00400018196562-0j)
Herbivorous true bugs	(0.0171771789136-0j)	Buffalo	(0.00400018196562-0j)
Millipedes	(0.0171406303957-0j)	Coqui francolin	(0.00396566262886-0j)
Fat or tree mouse	(0.0145964281845-0j)	Gray flycatcher	(0.00395357691141-0j)
Cape Dwarf Gecko	(0.0115137416235-0j)	Aardvark	(0.00393741953515-0j)
Dwarf Puddle Frog	(0.010450164574-0j)	Little Grebe	(0.00391433387594-0j)
Cape teal	(0.00878654125989-0j)	Four-toed African hedgehog	(0.00391433387594-0j)
Aquatic crustaceans, e.g. crabs	(0.00876977249064-0j)	Ground pangolin	(0.00383793981236-0j)
Yellow-bellied Eremomela	(0.0085803969266-0j)	Banana pipistrelle	(0.00381850124583-0j)
Cockroaches	(0.00838605224319-0j)	Yellow-throated Sandgrouse	(0.00379006426361-0j)
Dwarf Sand Snake	(0.00808066640798-0j)	Bat-eared fox	(0.00379006426361-0j)
Crops	(0.00795696094163-0j)	Nile monitor lizard	(0.00379006426361-0j)
Katydids	(0.0077535021439-0j)	Striped weasel	(0.00379006426361-0j)
Mueller's Clawed Frog	(0.00761623876969-0j)	Dwarf mongoose	(0.00379006426361-0j)
Rufous Beaked Snake	(0.00714536566069-0j)	Dwarf Epauletted Bat	(0.0037892200407-0j)
Shrew	(0.0067741364859-0j)	African wild dog	(0.00365498859803-0j)
Preying mantids	(0.0067062367942-0j)	Maasai giraffe	(0.00365498859803-0j)
Smith's red rock hare	(0.00648415744864-0j)	Lappet-faced vulture	(0.00365498859803-0j)
Thripidae Thrips	(0.00630714339135-0j)	African fish eagle	(0.00365498859803-0j)
Yellow-breasted apalis	(0.00629705257492-0j)	Black rhinoceros	(0.00365498859803-0j)
Three-banded plover	(0.00537221113923-0j)	Hooded vulture	(0.00365498859803-0j)
Creek rat	(0.00534259019486-0j)	Serval cat	(0.00365498859803-0j)
Scorpions	(0.00532137492359-0j)	African white-backed vulture	(0.00365498859803-0j)
White-bellied bustard	(0.00531141341877-0j)	Ostrich	(0.00365498859803-0j)
Centipedes	(0.00516931207869-0j)	Lion	(0.00365498859803-0j)
Red-faced Crombec	(0.00515209022569-0j)	Human	(0.00365498859803-0j)
Leopard tortoise	(0.00495266204991-0j)	Pearl-spotted owlet	(0.00365498859803-0j)
Topi	(0.0049000732511-0j)	Striped hyena	(0.00365498859803-0j)
Kirki's dikdik	(0.0047762381479-0j)	African marsh owl	(0.00365498859803-0j)
Bush hyrax	(0.00475039864177-0j)	Black backed jackal	(0.00365498859803-0j)

Bohor reedbuck	(0.00471712203677-0j)	African elephant	(0.00365498859803-0j)
Spot-necked otter	(0.00469212675355-0j)	Domestic cat	(0.00365498859803-0j)
Kori bustard	(0.00468145848158-0j)	Marabou stork	(0.00365498859803-0j)
Oribi	(0.00446556411707-0j)	Domestic dog	(0.00365498859803-0j)
Black-and-white colobus monkey	(0.00444056883385-0j)	Hippopotamus	(0.00365498859803-0j)
Greater bush baby	(0.00440752461197-0j)		

(d) Ranking by **closeness_centrality**:

Decaying material	1	Nile crocodile	0.393305439	Nile crocodile	0.393305439
Black backed jackal	0.576687117	African wild dog	0.391666667	African wild dog	0.391666667
Fat or tree mouse	0.528089888	African fish eagle	0.391666667	African fish eagle	0.391666667
Creek rat	0.525139665	White-bellied bustard	0.388429752	White-bellied bustard	0.388429752
Katydids	0.519337017	Marabou stork	0.386831276	Marabou stork	0.386831276
Rove- and ground beetles	0.519337017	Serval cat	0.385245902	Serval cat	0.385245902
Mueller's Clawed Frog	0.519337017	Lion	0.383673469	Lion	0.383673469
Bushpig	0.513661202	Little Grebe	0.383673469	Little Grebe	0.383673469
Crickets	0.510869565	Rufous Beaked Snake	0.382113821	Rufous Beaked Snake	0.382113821
Black-and-white colobus monkey	0.510869565	Greater bush baby	0.382113821	Greater bush baby	0.382113821
Formicidae Ants	0.508108108	Nile monitor lizard	0.382113821	Nile monitor lizard	0.382113821
Harvester termites	0.505376344	Yellow-bellied Eremomela	0.380566802	Yellow-bellied Eremomela	0.380566802
Millipedes	0.505376344	African white-backed vulture	0.380566802	African white-backed vulture	0.380566802
Thripidae Thrips	0.505376344	African Grey hornbill	0.380566802	African Grey hornbill	0.380566802
Aquatic crustaceans, e.g. crabs	0.505376344	Lappet-faced vulture	0.37751004	Lappet-faced vulture	0.37751004
Flies	0.505376344	African clawless otter	0.376	African clawless otter	0.376
Grass and herbs	0.502673797	Leopard tortoise	0.376	Leopard tortoise	0.376
Crops	0.502673797	Dwarf Sand Snake	0.373015873	Dwarf Sand Snake	0.373015873
Fruits and nectar	0.502673797	Yellow-breasted apalis	0.373015873	Yellow-breasted apalis	0.373015873
Trees and shrubs	0.502673797	Striped weasel	0.3671875	Striped weasel	0.3671875
Grains, seeds	0.502673797	Four-toed African hedgehog	0.365758755	Four-toed African hedgehog	0.365758755
Cockroaches	0.502673797	Three-banded plover	0.364341085	Three-banded plover	0.364341085
Plant juices	0.502673797	Striped hyena	0.362934363	Striped hyena	0.362934363

Fish, e.g. catfish	0.502673797	Centipedes	0.362934363	Centipedes	0.362934363
African marsh owl	0.45410628	Human	0.361538462	Human	0.361538462
Leopard	0.425339367	Shrew	0.361538462	Shrew	0.361538462
Dwarf mongoose	0.425339367	Dwarf Puddle Frog	0.360153257	Dwarf Puddle Frog	0.360153257
Pearl-spotted owlet	0.408695652	Gray flycatcher	0.358778626	Gray flycatcher	0.358778626
Hooded vulture	0.406926407	Cape Dwarf Gecko	0.356060606	Cape Dwarf Gecko	0.356060606
Southern ground hornbill	0.406926407	Orb-weaver spiders	0.356060606	Orb-weaver spiders	0.356060606
Black mamba	0.403433476	Cape teal	0.353383459	Cape teal	0.353383459
Bat-eared fox	0.398305085	Preying mantids	0.352059925	Preying mantids	0.352059925

3. Appendix C: **Coding**

```
import numpy as np
import copy
from scipy import linalg as la
from collections import OrderedDict, defaultdict
from operator import itemgetter
import csv
import networkx as nx
import matplotlib.pyplot as plt

# read both predator-prey relation file and label file
# construct a original graph as well as a modified graph

def readFile():
    graph = defaultdict(list)
    label = defaultdict(list)
    reversedGraph = defaultdict(list)
    with open("Predator-Prey Relations.txt") as f1:
        for line in f1:
            (predator, prey) = line.split(';')
            graph[int(prey)].append(int(predator))
            reversedGraph[int(predator)].append(int(prey))
    with open("Label.txt") as f2:
        for line in f2:
            (species, num) = line.split(';')
            label[int(num)] = species

    # get the original graph by adding unconnected vertices
    for key in label:
        graph[key]
    # deep copy the graph as original graph
    originalGraph = copy.deepcopy(graph)

    # get the modified graph by
    # adding edges from every node to decaying material
    # get the reversed graph by
    # adding edges frim decaying material to all other nodes
    # and adding edges from unconnected vertices to decaying material
    for key in label:
```

```
        if key != 1:
            graph[key].append(1)
            reversedGraph[1].append(key)

    return originalGraph, graph, reversedGraph, label

# plot given graph with labels
def plot(G, label):
    # G=nx.DiGraph(graph)
    newLabel = dict()
    # truncate the label
    for num in label:
        data = label[num]
        if len(data) > 15:
            newLabel[num] = (data[:15] + '...')
        else:
            newLabel[num] = data
    # replace labels of numbers with labels of names
    H=nx.relabel_nodes(G,label)
    centrality = nx.eigenvector_centrality_numpy(H)
    print(['%s %0.2f'%(node,centrality[node]) for node in centrality])

    # set size of canvas
    plt.figure(figsize=(14,10))
    # pos = nx.random_layout(H) # random layout
    pos = nx.spring_layout(H,k=1,iterations=30) # Force-directed graph drawing
    nx.draw(H,pos,node_size=500,font_size=9,node_color='#A0CBE2',edge_color='#BB0000',width=0.5,edge_cmap=plt.cm.Blues,arrows=True,with_labels=True)
    plt.show()

# get the limiting distribution of given transition matrix
# return the rank of node by ranking the limiting distribution
def rank(G, label):
    # The transition matrix of a Markov Chain always has an eigenvalue 1
    # find the eigenvector corresponding to the eigenvalue 1
    e_vals, e_vecs = la.eig(G.T)
    i = 0
    while np.abs(e_vals[i] - 1.) > 1e-8:
        i = i + 1
    limiting_distribution = e_vecs[:, i]
    # normalize the eigenvector we found
    limiting_distribution = limiting_distribution / np.sum(limiting_distribution)
    d = dict()
    for i in label:
        d[label[i]] = limiting_distribution[i - 1]
    rank = OrderedDict(sorted(d.items(), key=itemgetter(1), reverse=True))
    return rank

# create a markov chain based on the given graph
# return a transition matrix of the markov chain
def get_markov_chain(graph, N):
    transition_matrix = np.zeros(shape = (N, N))
    for i in range(N):
        for j in range(N):
            edges = graph[i+1]
            if j+1 in edges:
                degree = len(edges)
```

```
                    transition_matrix[i][j] = 1. / degree
        return transition_matrix

    # construct a new matrix with damping factor
    # by using pagerank algorithm
    def get_google_matrix(T, N):
        d = 0.85 # damping factor
        G = d * T + (1-d)/N * np.ones((N, N))
        return G

    # return a rank by degree of vertices
    def rankByDegree(graph, label):
        d = dict()
        for i in label:
            d[label[i]] = len(graph[i])
        rank = OrderedDict(sorted(d.items(), key=itemgetter(1), reverse=True))
        return rank

    # return a rank by average shortest path
    def rankByCloseness(G, label):
        centrality = nx.closeness_centrality(G)
        d = dict()
        for i in label:
            d[label[i]] = centrality[i]
        rank = OrderedDict(sorted(d.items(), key=itemgetter(1), reverse=True))
        return rank

    # output the rank as a csv file
    # (the csv file has already created, we only modified the file)
    def output(dict1, filename):
        with open(filename, 'wb') as output:
            writer = csv.writer(output)
            for key, value in dict1.iteritems():
                writer.writerow([key, value])

    if __name__ == '__main__':
        originalGraph, graph, reversedGraph, label = readFile()
        N = len(label)
        TT = get_markov_chain(graph, N)
        T = get_markov_chain(reversedGraph, N)
        G = get_google_matrix(T, N)
        rank1 = rankByDegree(originalGraph, label)
        rank2 = rank(T, label)
        rank3 = rank(G, label)
        rank4 = rank(TT, label)
        output(rank4,'output.csv')
        G_nx=nx.DiGraph(reversedGraph)
        rank4 = rankByCloseness(G_nx, label)
        output(rank1, 'output1.csv')
        output(rank2, 'output2.csv')
        output(rank3, 'output3.csv')
        output(rank4, 'output4.csv')
        # plot(G_nx, label)
```