

# Analytická úloha

## Dataset

Pre analytickú úlohu som vybral sieť **ARXIV NETSCIENCE MULTIPLEX**, ktorá reprezentuje spoluautorstvo. Vychádza z dát elektronického archívu preprintu vedeckých prác z oblasti matematiky, fyziky, informatiky a biológie.

## Základné štatistiky

Počet vrcholov	<b>14489</b>
Počet hrán	<b>59026</b>

## Základné štatistiky viacvrstvového grafu

Názov	Hodnota	Actor	Vrstva/vrstvy
Najvyšší stupeň	<b>361</b>	172	1,2,3,5,6,7,8,12
Najvyššie susedstvo	<b>151</b>	172	2,3,5,6,8
Najvyššia redundancia	<b>0.833</b>	3747	1,2,3,6,8,12
Najvyššie exkluzívne susedstvo	<b>151</b>	172	1,2,3,5,6,7,8,12

Sieť je rozdelená na 13 vrstiev, podľa jednotlivých kategórií:

1. **physics.soc-ph**
2. **physics.data-an**
3. **physics.bio-ph**
4. math-ph
5. math.OC
6. cond-mat.dis-nn
7. cond-mat.stat-mech
8. q-bio.MN
9. q-bio
10. q-bio.BM
11. nlin.AO
12. cs.SI
13. cs.CV

Pre ďalšiu analytickú analýzu boli zvolené 3 vrstvy z rovnakej kategórie – fyziky, ktoré sú vyznačené v zozname.

Počet vrcholov: **1594**

Počet hrán: **3019**

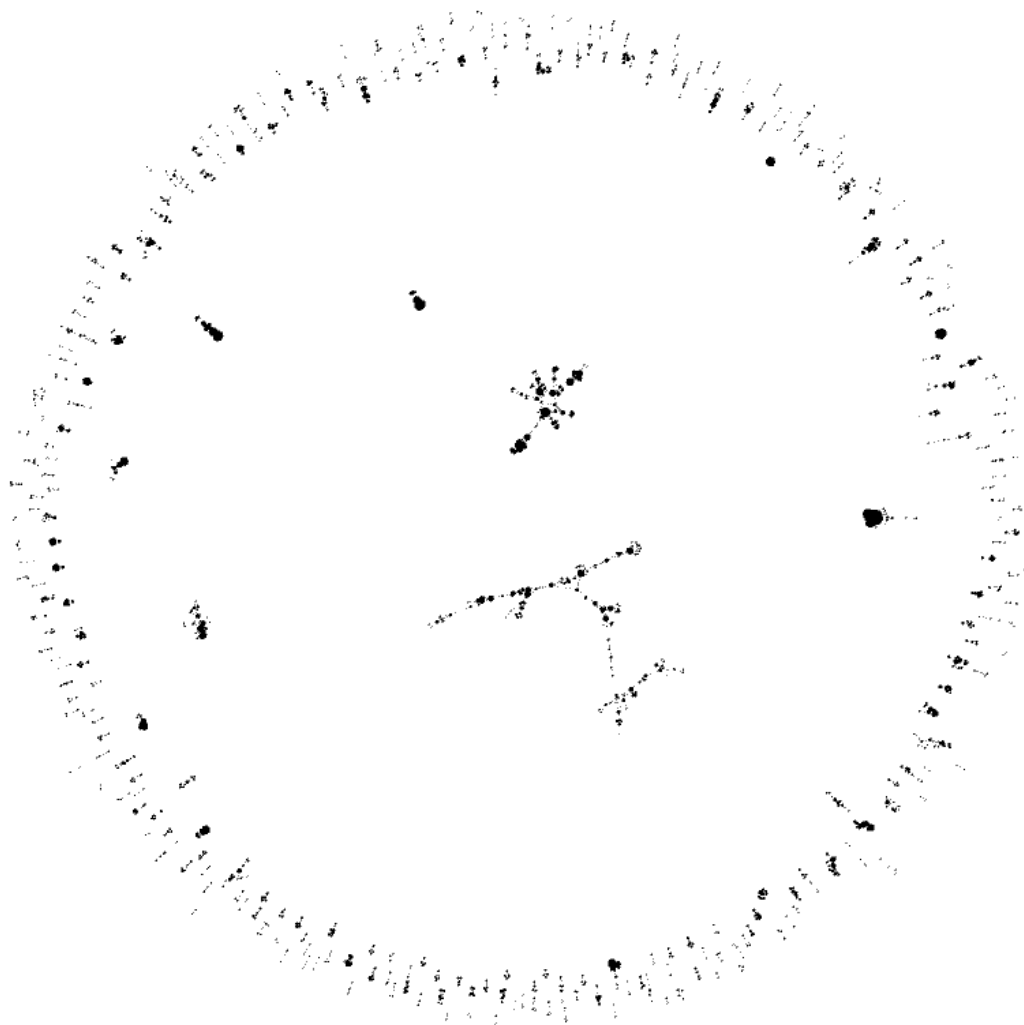
Hustota: **0.0024**

Priemerný stupeň: **3.788**

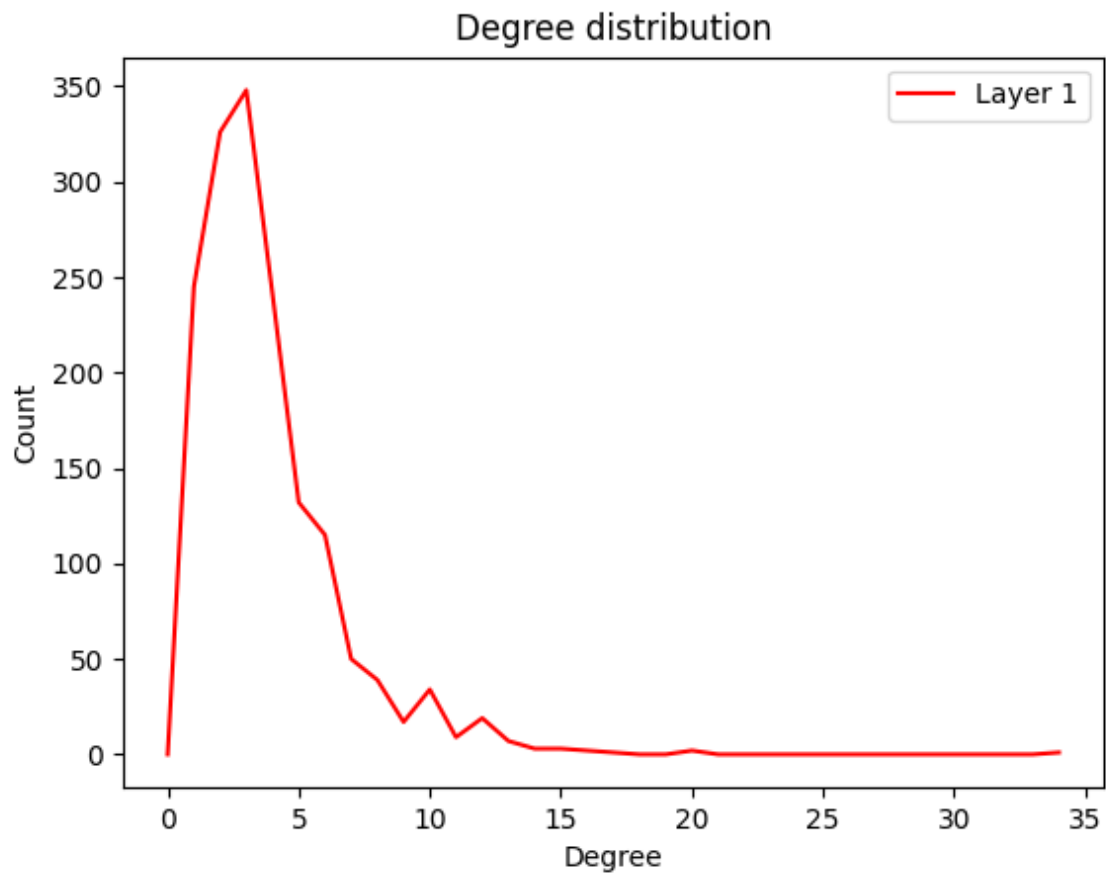
Maximálny stupeň: **34**

Priemerný zhlukovací koeficient: **0.775**

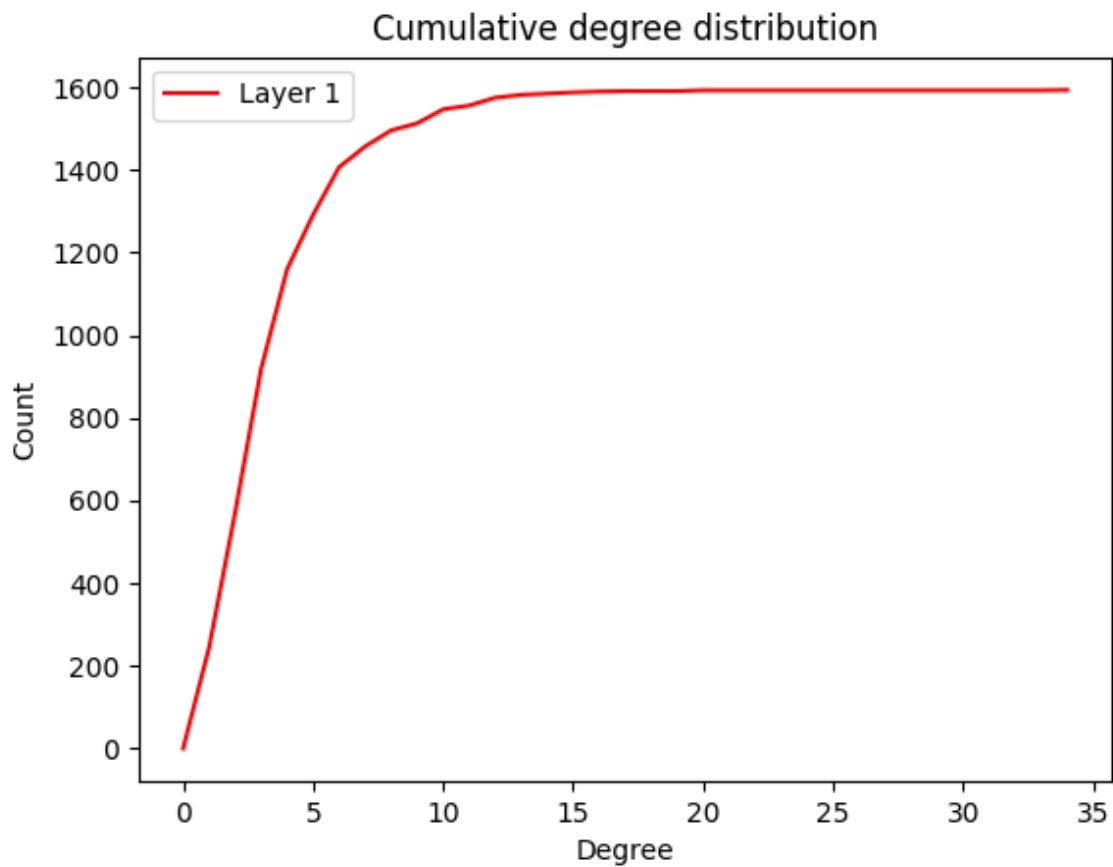
Počet komponent: **310**



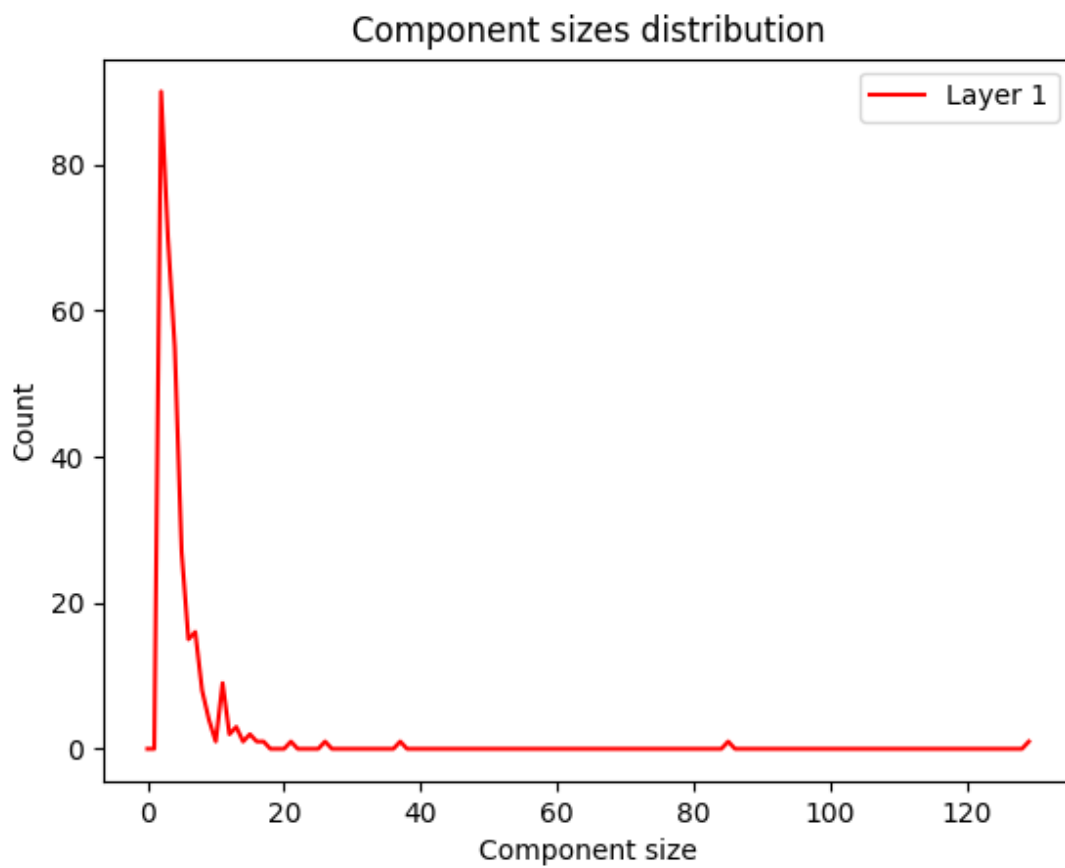
Obrázok 1 - Vizualizácia



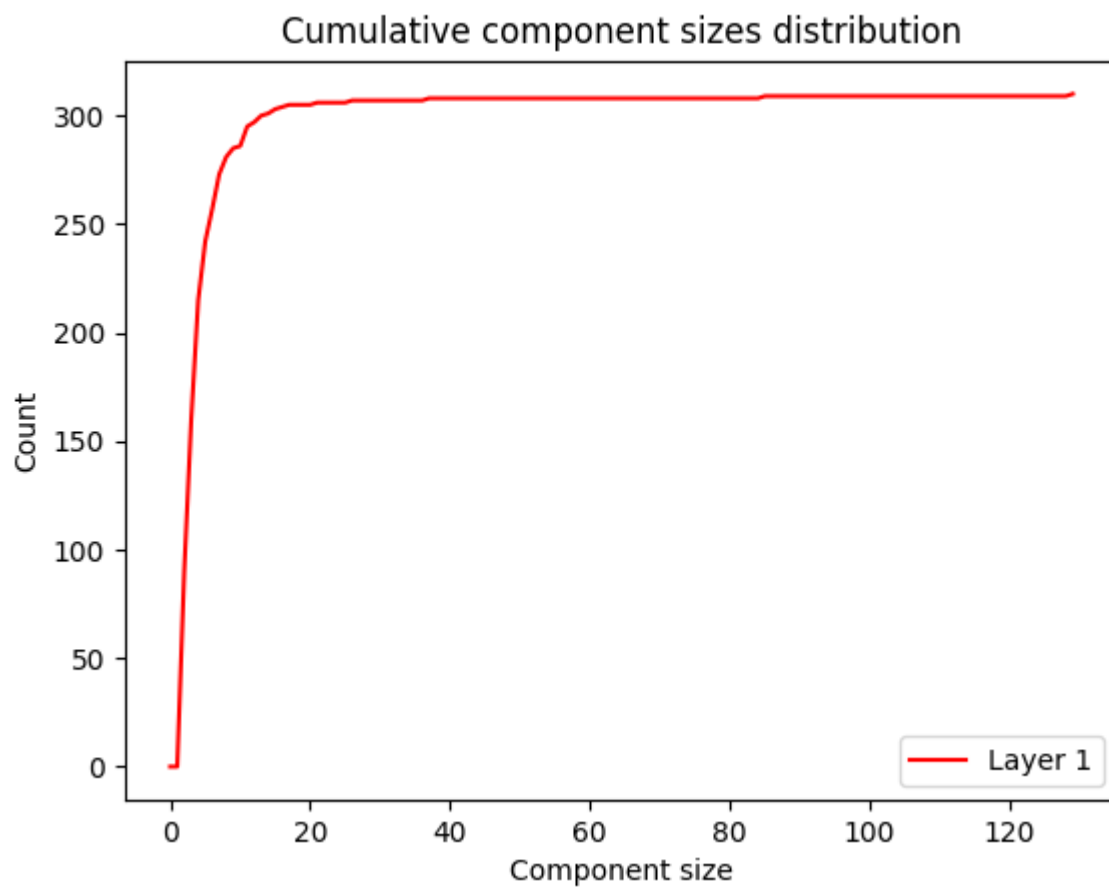
Obrázok 2 - Distribúcia stupňov



Obrázok 3 - Kumulatívna distribúcia stupňov



Obrázok 4 - Distribúcia veľkostí komponent



Obrázok 5 - Kumulatívna distribúcia veľkostí stupňov

Počet vrcholov: **5465**

Počet hrán: **14485**

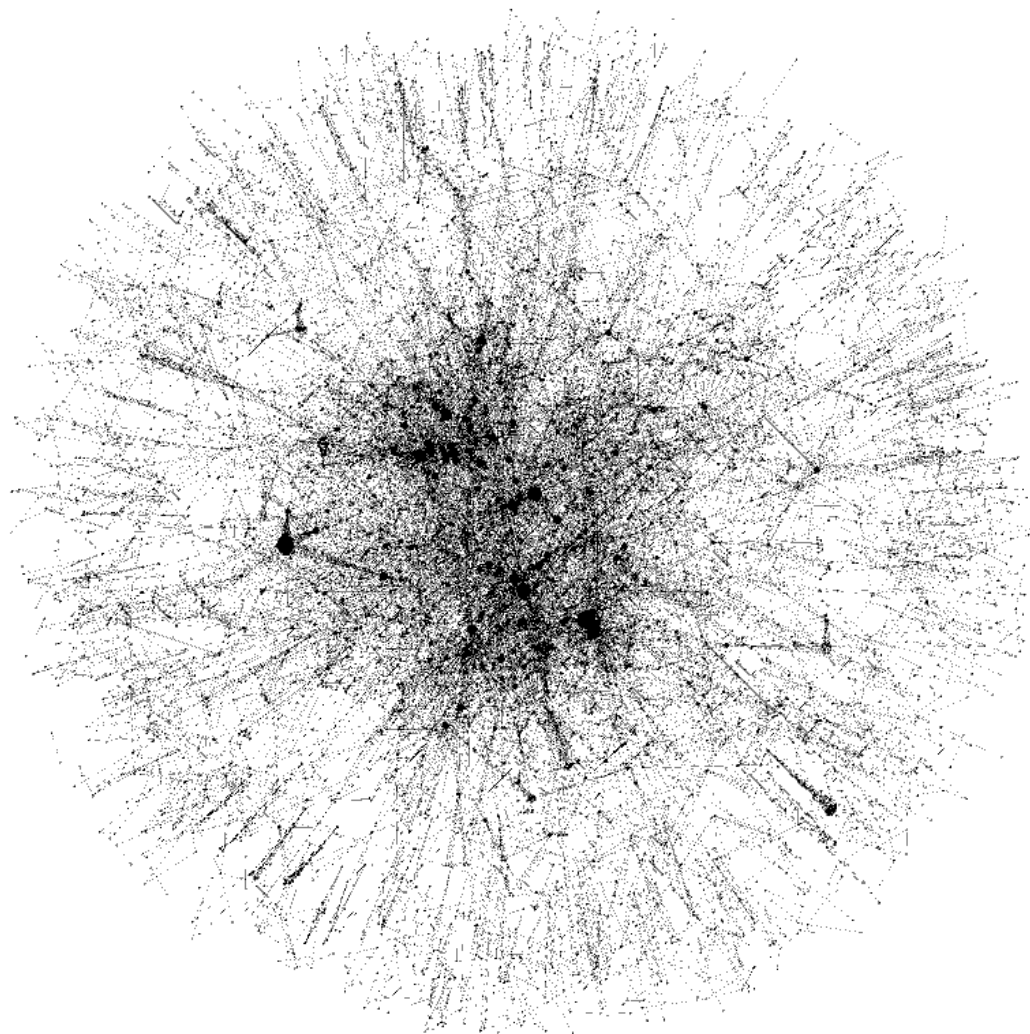
Hustota: **0.00097**

Priemerný stupeň: **5.30**

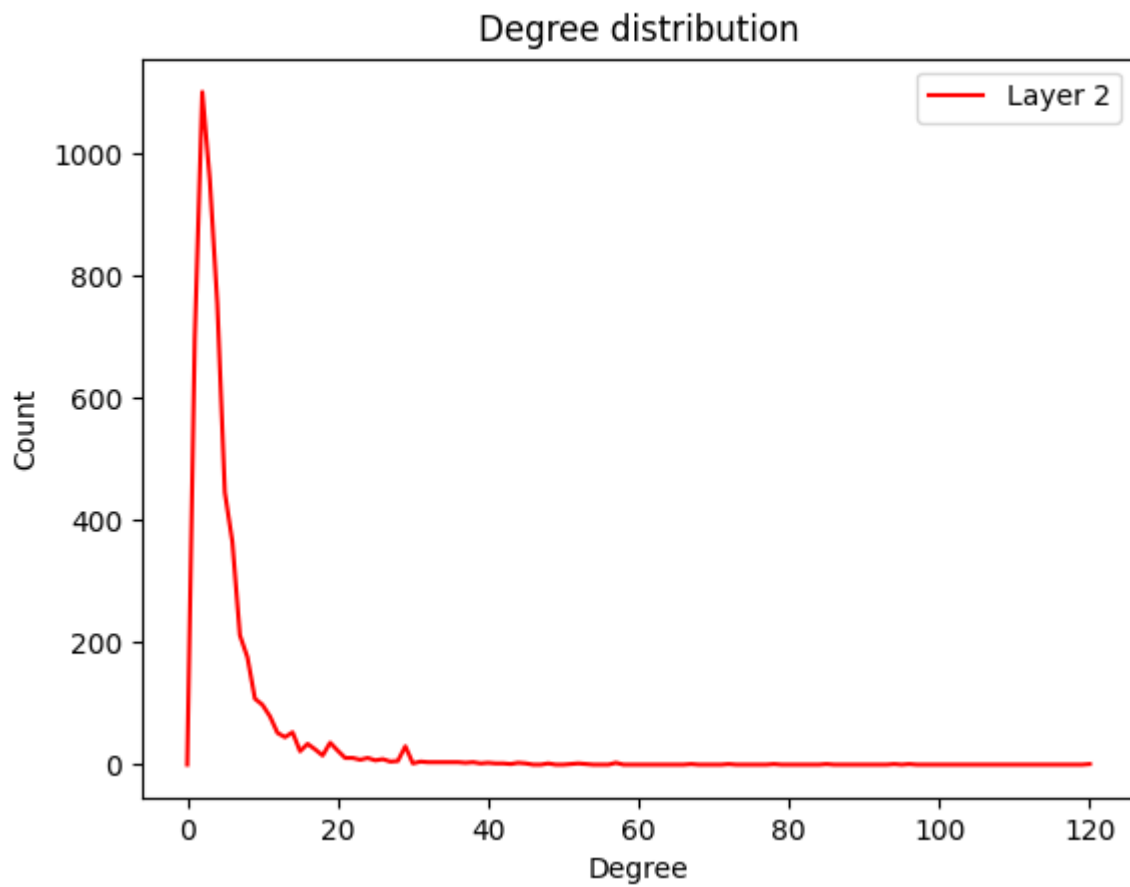
Maximálny stupeň: **120**

Priemerný zhukovací koeficient: **0.72**

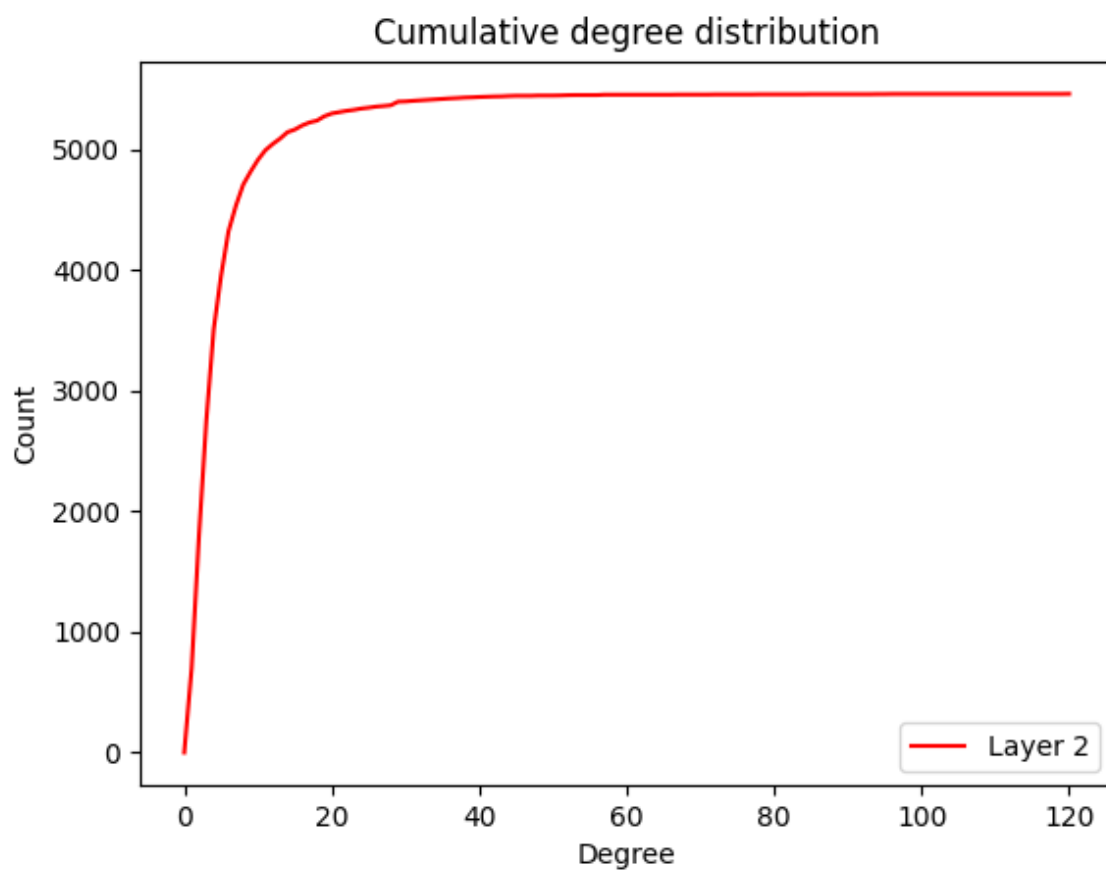
Počet komponent: **508**



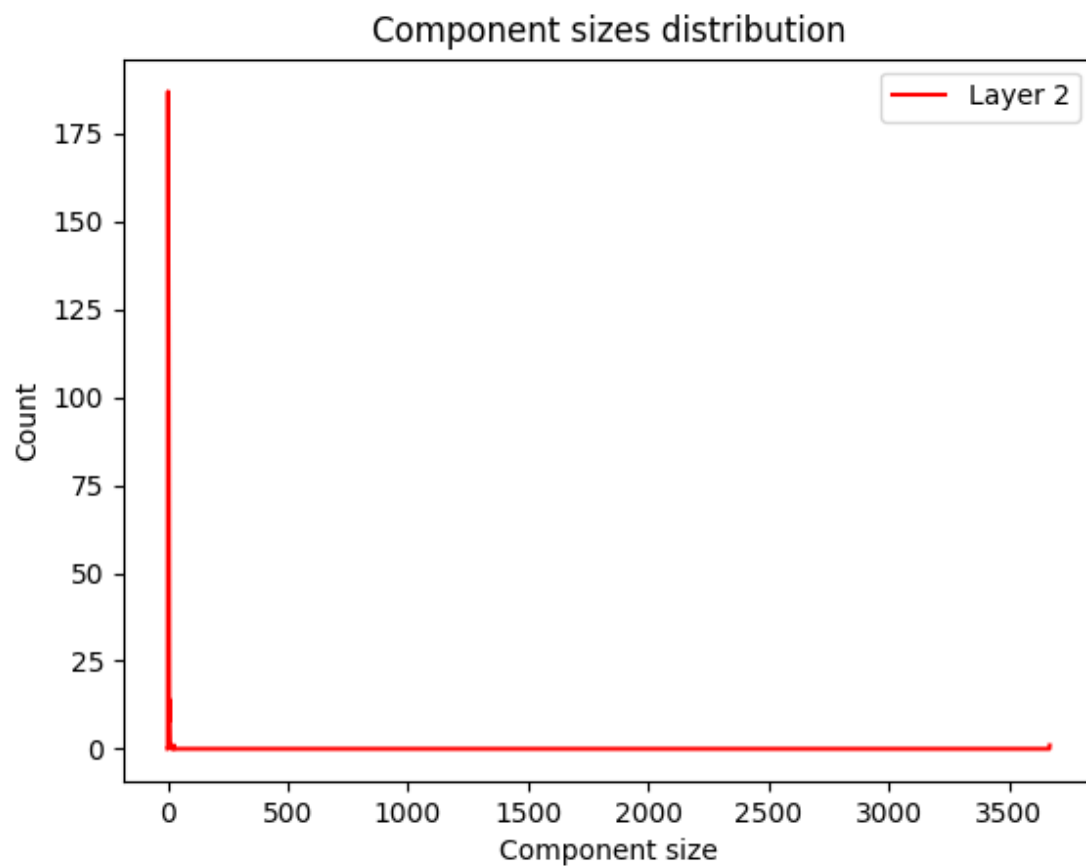
Obrázok 6 - Vizualizácia



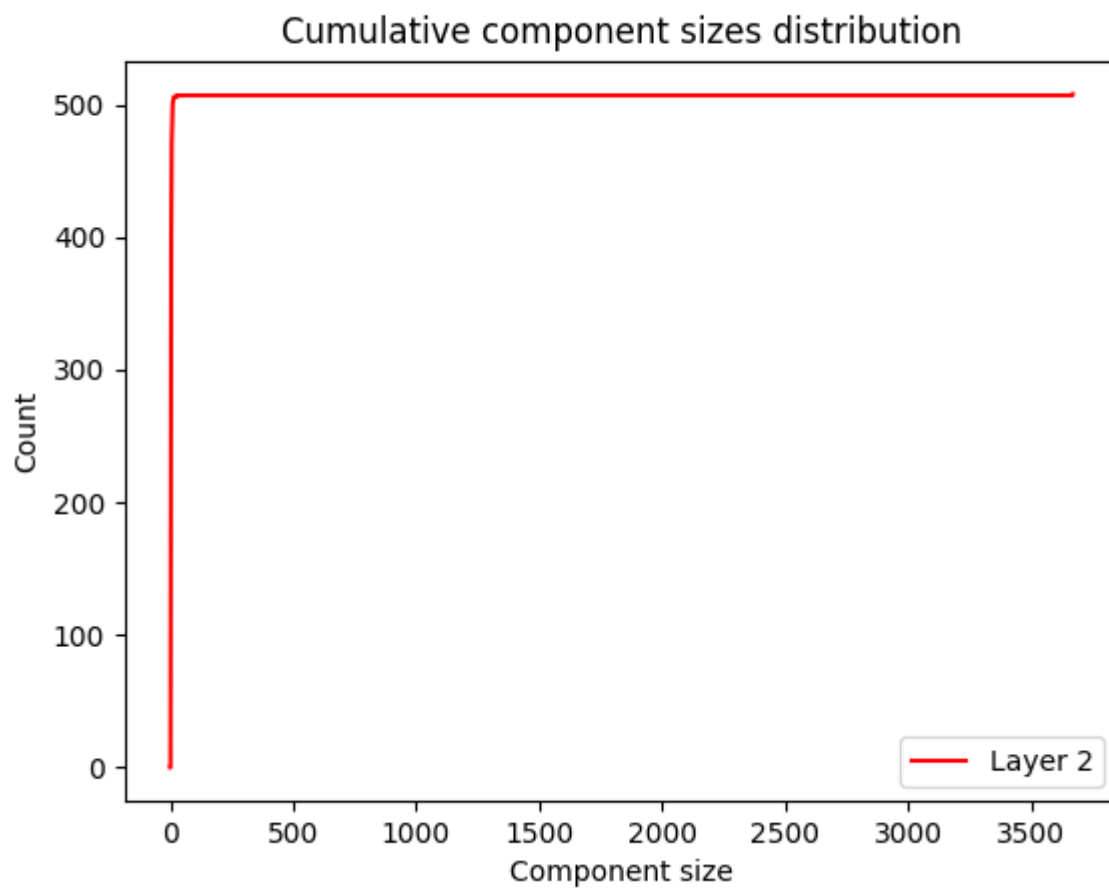
Obrázok 7 - Distribúcia stupňov



Obrázok 8 - Kumulatívna distribúcia stupňov



Obrázok 9 - Distribúcia veľkostí komponent



Obrázok 10 - Kumulatívna distribúcia veľkostí stupňov

### Vrstva 3 – physics.bio-ph

Počet vrcholov: **2956**

Počet hrán: **6097.0**

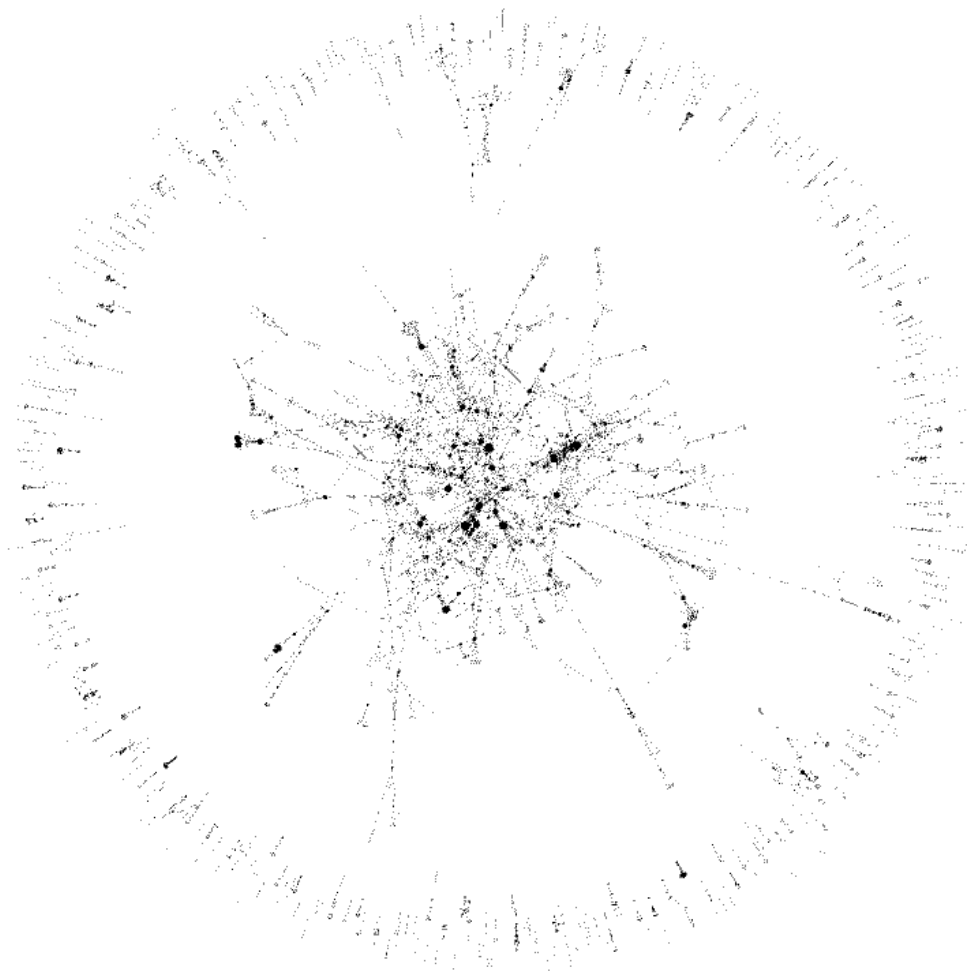
Hustota: **0.0014**

Priemerný stupeň: **4.125**

Maximálny stupeň: **46**

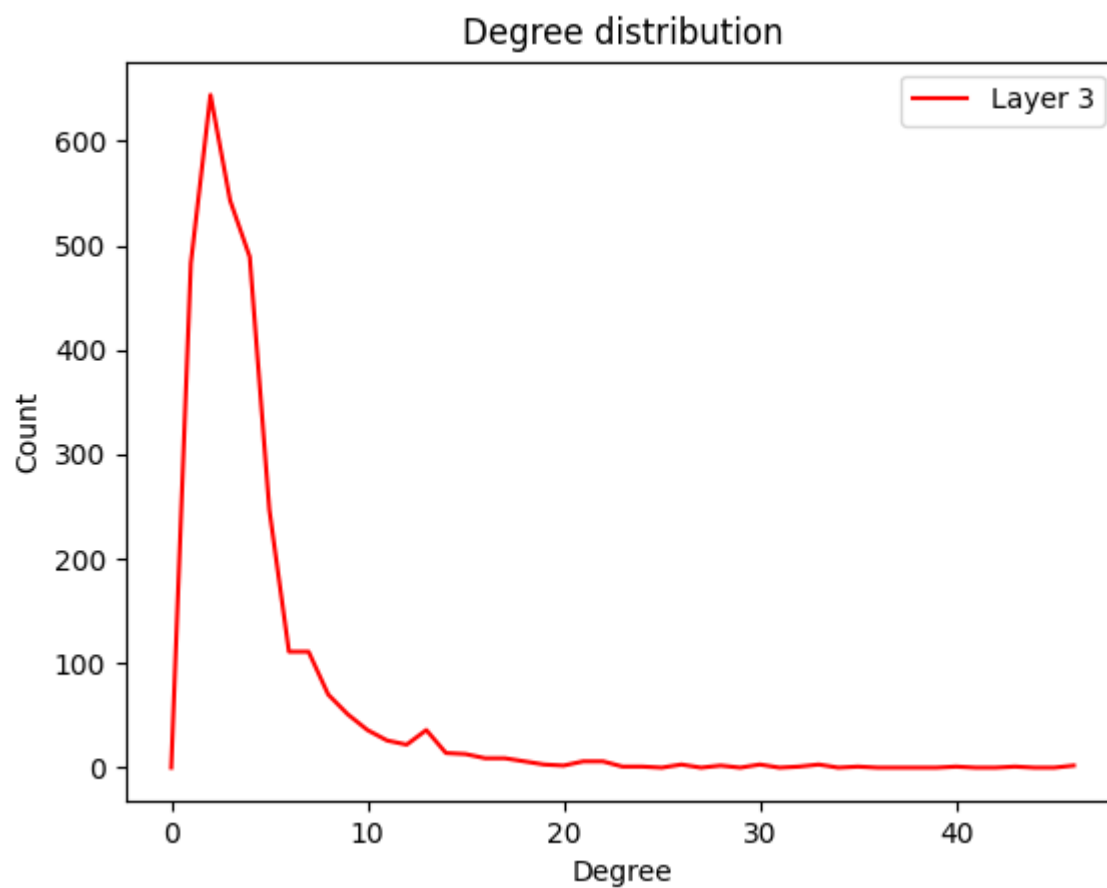
Priemerný zhukovací koeficient: **0.696**

Počet komponent: **348**

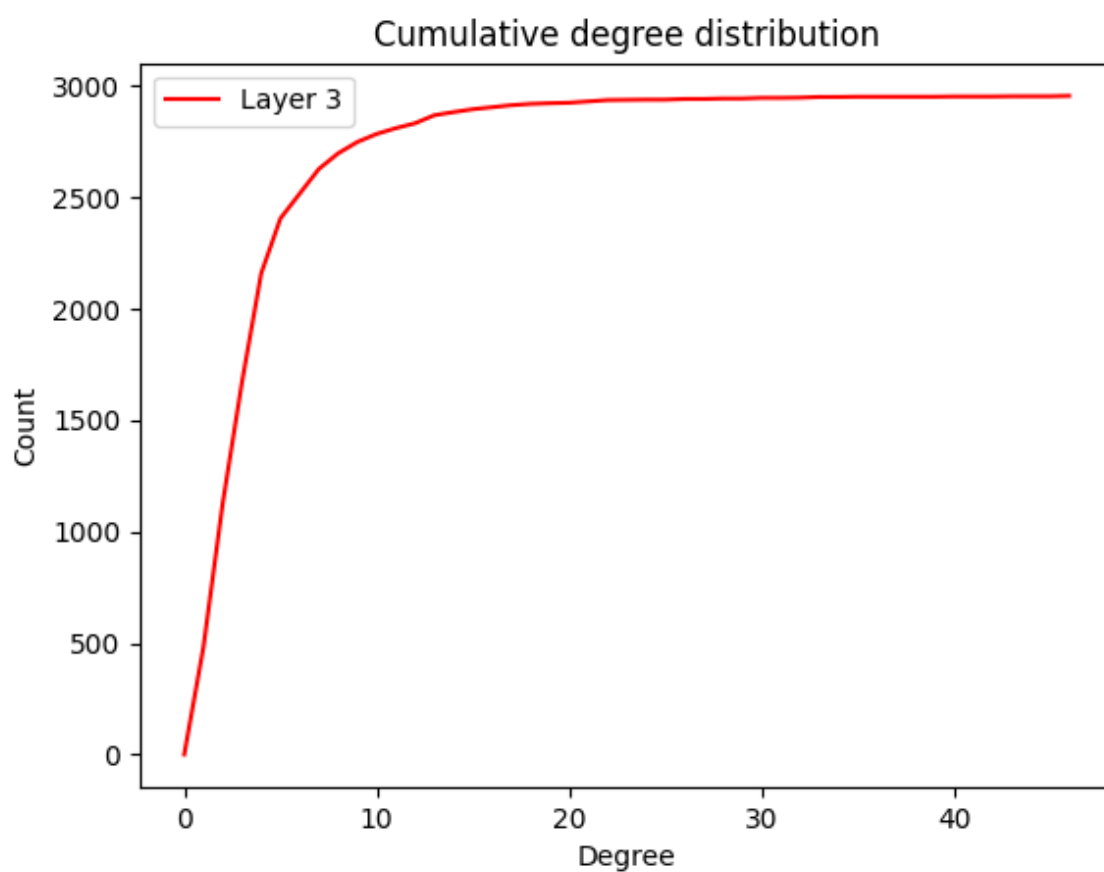


Obrázok 11 - Vizualizácia

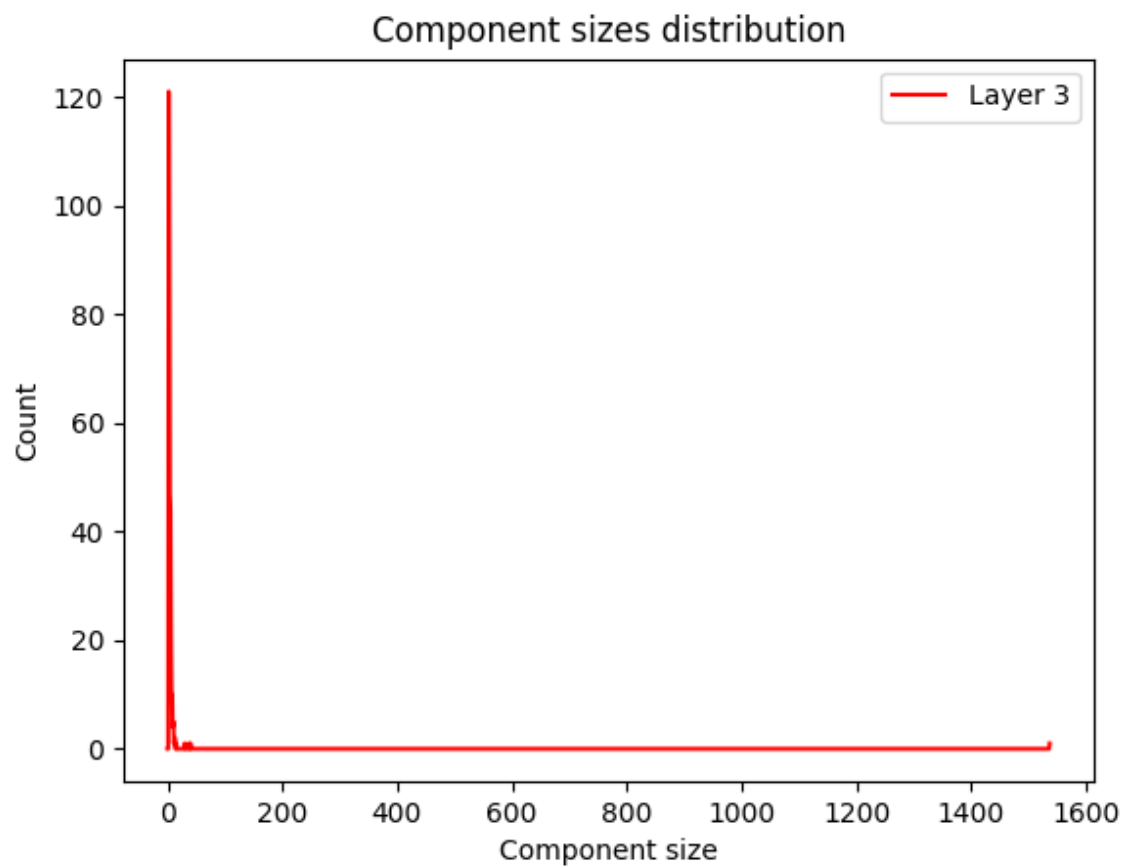




Obrázok 12 - Distribúcia stupňov



Obrázok 13 - Kumulatívna distribúcia stupňov



## Záver

Na základe vizualizácií a analýzy veľkosti komponent môžeme usúdiť, že veľkosť komponenty, teda počet spolupracujúcich autorov, je silno závislý na tematike danej práce – vo vrstve 1 vidíme najmä menšie komponenty, zatiaľ čo vrstvy 2 a 3 obsahujú ako malé komponenty, tak jednu veľkú, ktorá dominuje celému grafu.

V 2. vrstve zároveň vidíme vysoký počet vrcholov so stupňom 1, čo znamená, že majú len jedného suseda. Tento jav nepozorujeme na ostatných vrstvách, kde je priemerný stupeň vrcholov 2, resp. 3.