Evolution of academic knowledge networks in transition economies – the Chinese perspective

© Victoria Vesna

Structure:

- The Theory
- The Policy
- The Methodology
- Method of Calculation
- The Data
- The Results

Stefan Hennemann (Gießen, Germany)
June 27th 2008
Innovation Systems, Knowledge Creation and Knowledge Networks

Company R&D combine knowledge stemming from different sources

Public research organizations (PRO) and higher education institutions (HEI) are important sources of knowledge

Some excellent PROHEI in developing countries acquire knowledge from foreign equivalent organizations and adapt it to the local context

With further development other sub-centers develop and build up a hierarchical system of foreign knowledge acquisition and local dissemination

The metaphor of a hierarchical knowledge distribution system is discussed in the innovation systems and innovation networks literature
The Policy

- Complete reorganization of former soviet-style knowledge creation system from the mid 1980s onwards
- PRO as well as HEI consolidated and combined to form capable acting organizations (1990s)
- Identification of some 100 universities of excellence (1990s, 211-project) to form internationally acknowledged and competitive organizations
- Reorder of responsibilities for the HEI:
  (a) state level, central education and resort ministry attached
  (b) provincial level, provincial or education ministry attached
  (c) local government attached
Global Universities
e. g. Harvard Univ, Oxford Univ, Univ Singapore, Kyoto Univ, CALTECH, MIT

State Level 211-Universities (a)
e. g. Peking Univ, Tsing Hua Univ, Fudan Univ, Harbin TECH

Provincial Level Universities (b)
e. g. Zhongshan Univ, HK Polytech, Sichuan Union Univ

Hypotheses:
(1) The reorganization of the S&T-System supports the development of a hierarchical structure of knowledge-networks both intra-regional and inter-regional.
(2) T1-Universities take an interlink / mediating position between the regional and the global sphere
(3) The T2 integration into the global network improves over time
The Methodology (I)

- Bibliometrical information and co-authorship is largely available from citation databases (problematic: poor quality of the raw data)
  - Database: Thompson SCI/Expanded, SSCI

- Pre-assumptions/constraints:
  - At least one author of a publication has to be registered with a Chinese organization
  - Classification and grouping by organizations not by individual authors
  - intra-organizational flows are optimized
  - relevant knowledge flows between co-authors are symmetrical
  - knowledge flows have a max. reach of 2 edges
  - Multiple collaborations between two organizations account for higher relevance
  - more than 10 co-authors are not able to share knowledge comprehensively
  - there are no self-loops (different authors of the same organization)
The Methodology (II)

- Timeseries from 1997 to 2007
- Tier-0 criteria: combined global university rankings of *The Times* and *newsweek* both based on expert opinions (T0)
- Tier-1 criteria: 211-project Universities (T1)
- Tier-2 criteria: rest of Chinese HEI (T2)
- Given specifications lead to a multiedge, unipartite and undirected network
- Identification of largest subnetwork without disconnected parts (subnet has scale-free properties $p(k) \sim k^{-2.0}$)
- Calculation of cross-tier connectivity within two steps
Method of Calculation (I)

Direct connectivity from one tier to another (degree)

- T1 nodes reach on average 1.5 T0 nodes directly
- 1.5 T0 nodes represent 18.75% of all T0 nodes
Indirect connectivity (2nd order reachability) from one tier to another

- T1 nodes reach on average 3 T0 nodes within two steps
- 3 T0 nodes represent 37.5% of all T0 nodes
### Basic Statistics for the Graph/Network

<table>
<thead>
<tr>
<th>Year</th>
<th>Size</th>
<th>Density</th>
<th>Edges</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.457</td>
<td>0.169</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>1998</td>
<td>0.519</td>
<td>0.219</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>1999</td>
<td>0.490</td>
<td>0.221</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>2000</td>
<td>0.508</td>
<td>0.228</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>2001</td>
<td>0.428</td>
<td>0.233</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>2002</td>
<td>0.526</td>
<td>0.242</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>2003</td>
<td>0.436</td>
<td>0.235</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>2004</td>
<td>0.533</td>
<td>0.238</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>2005</td>
<td>0.586</td>
<td>0.256</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>2006</td>
<td>0.590</td>
<td>0.261</td>
<td>22</td>
<td>7</td>
</tr>
<tr>
<td>2007</td>
<td>0.606</td>
<td>0.261</td>
<td>23</td>
<td>7</td>
</tr>
</tbody>
</table>

Note: all calculations were done with the largest complete subgraph rather than with all available nodes in the respective year.
Connectivity and Hierarchy to Tier-0 Organizations

- T1-Nodes are increasingly directly connected to T0-Nodes.
- T2-Nodes are weakly connected directly to T0 and cannot improve their situation over time.
- The 2nd order connectivity of T1 to T0 is remarkably stable. They reach nearly half of T0-Nodes within two edges. T2 connectivity increases in this respect which might be a hint for better connectedness to T1.
Connectivity and Hierarchy to Tier-1 Organizations

- T1-Nodes connect better to other T1-Nodes than to other tiers (rather stable ~7-8%)
- Surprisingly, direct connection of T2 to T1 is low and even decreases slightly over time
- The 2nd order connectedness of T2 to T1 improves in the first years and stays at high level
Connectivity and Hierarchy to Tier-2 Organizations

Direct connections to higher levels T0 and T1 is decreasing or at best stable, and so is the intra-tier-connectedness.

Thus, the improving 2nd order connectedness of T0 and T1 can neither be attributed to higher local level nor to higher global levels.
The Chinese knowledge network is evolving into a better connected one over time and the reorganization of the Chinese S&T-System is obviously successful.

Chinese elite universities (T1) have a strong and increasingly dominant position to connect China to the global knowledge system.

The improvement of the indirect connectedness of T2 to the global leaders T0 can be ascribed to an improving direct link to foreign non-T0 organizations and NOT to mediating T1.

This result has two direct implications: (1) the hierarchy hypotheses have to be partly rejected, (2) a „club“-hypothesis can be proposed.

Side-effect of the research: The stability over the years and the relatively stable degree distribution proves the theoretical considerations and simulation to scale-free networks (cp. Albert and Barabási 2002).
Further Research | Method-Tuning

- Bipartite graph based on identified cliques of organizations?
- Different method of calculation?
- Stronger regional differentiation or differentiation of academic fields; company focus; introducing additional countries (benchmarking)
- Analysis of the most interdisciplinary actors
- General question of random sampling and effects on the results
I would like to thank Dr. Diego Rybski (New York), Prof. Eike W. Schamp (Frankfurt) and Prof. Ingo Liefner as well as other colleagues from Giessen University for their helpful comments and discussion!
The data have been binned logarithmically to reduce noise.
### Mean-Scores

<table>
<thead>
<tr>
<th>Year</th>
<th>T0</th>
<th>T1</th>
<th>T2</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.00302</td>
<td>0.01509</td>
<td>0.00206</td>
<td>0.00133</td>
<td>0.01153</td>
<td>0.00060</td>
<td>1997</td>
</tr>
<tr>
<td>1998</td>
<td>0.00386</td>
<td>0.01929</td>
<td>0.00188</td>
<td>0.00090</td>
<td>0.00931</td>
<td>0.00050</td>
<td>1998</td>
</tr>
<tr>
<td>1999</td>
<td>0.00315</td>
<td>0.01957</td>
<td>0.00157</td>
<td>0.00079</td>
<td>0.00711</td>
<td>0.00045</td>
<td>1999</td>
</tr>
<tr>
<td>2000</td>
<td>0.00313</td>
<td>0.02159</td>
<td>0.00156</td>
<td>0.00066</td>
<td>0.00704</td>
<td>0.00040</td>
<td>2000</td>
</tr>
<tr>
<td>2001</td>
<td>0.00599</td>
<td>0.04201</td>
<td>0.00276</td>
<td>0.00119</td>
<td>0.00908</td>
<td>0.00068</td>
<td>2001</td>
</tr>
<tr>
<td>2002</td>
<td>0.00319</td>
<td>0.02144</td>
<td>0.00143</td>
<td>0.00058</td>
<td>0.00557</td>
<td>0.00037</td>
<td>2002</td>
</tr>
<tr>
<td>2003</td>
<td>0.00539</td>
<td>0.03830</td>
<td>0.00235</td>
<td>0.00094</td>
<td>0.00753</td>
<td>0.00053</td>
<td>2003</td>
</tr>
<tr>
<td>2004</td>
<td>0.00294</td>
<td>0.02065</td>
<td>0.00143</td>
<td>0.00056</td>
<td>0.00483</td>
<td>0.00031</td>
<td>2004</td>
</tr>
<tr>
<td>2005</td>
<td>0.00372</td>
<td>0.02160</td>
<td>0.00139</td>
<td>0.00051</td>
<td>0.00506</td>
<td>0.00027</td>
<td>2005</td>
</tr>
<tr>
<td>2006</td>
<td>0.00338</td>
<td>0.02064</td>
<td>0.00133</td>
<td>0.00047</td>
<td>0.00459</td>
<td>0.00025</td>
<td>2006</td>
</tr>
<tr>
<td>2007</td>
<td>0.00367</td>
<td>0.01820</td>
<td>0.00116</td>
<td>0.00040</td>
<td>0.00415</td>
<td>0.00021</td>
<td>2007</td>
</tr>
</tbody>
</table>

- General decline in the period of analysis for non HEIs and the Tier-2 HEIs
- Stable or slight increase of local (e.g. local in terms of network) relevance of T0 and T1
### Mean-Scores

Overall closeness increasing hint for better overall connection

Scores for Tier-2 HEIs increase slightly faster than for Tier-1 HEIs, the same holds true for company and T0

<table>
<thead>
<tr>
<th>Year</th>
<th>Tier-1</th>
<th>Tier-2</th>
<th>Tier-3</th>
<th>Tier-4</th>
<th>Tier-5</th>
<th>Tier-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.290</td>
<td>0.315</td>
<td>0.266</td>
<td>0.264</td>
<td>0.291</td>
<td>0.257</td>
</tr>
<tr>
<td>1998</td>
<td>0.319</td>
<td>0.335</td>
<td>0.288</td>
<td>0.277</td>
<td>0.306</td>
<td>0.274</td>
</tr>
<tr>
<td>1999</td>
<td>0.314</td>
<td>0.340</td>
<td>0.286</td>
<td>0.281</td>
<td>0.297</td>
<td>0.278</td>
</tr>
<tr>
<td>2000</td>
<td>0.323</td>
<td>0.347</td>
<td>0.294</td>
<td>0.286</td>
<td>0.308</td>
<td>0.285</td>
</tr>
<tr>
<td>2001</td>
<td>0.327</td>
<td>0.351</td>
<td>0.299</td>
<td>0.292</td>
<td>0.312</td>
<td>0.289</td>
</tr>
<tr>
<td>2002</td>
<td>0.327</td>
<td>0.346</td>
<td>0.301</td>
<td>0.294</td>
<td>0.311</td>
<td>0.292</td>
</tr>
<tr>
<td>2003</td>
<td>0.331</td>
<td>0.352</td>
<td>0.306</td>
<td>0.299</td>
<td>0.315</td>
<td>0.296</td>
</tr>
<tr>
<td>2004</td>
<td>0.334</td>
<td>0.352</td>
<td>0.310</td>
<td>0.303</td>
<td>0.314</td>
<td>0.302</td>
</tr>
<tr>
<td>2005</td>
<td>0.345</td>
<td>0.358</td>
<td>0.317</td>
<td>0.310</td>
<td>0.324</td>
<td>0.308</td>
</tr>
<tr>
<td>2006</td>
<td>0.342</td>
<td>0.359</td>
<td>0.317</td>
<td>0.312</td>
<td>0.322</td>
<td>0.308</td>
</tr>
<tr>
<td>2007</td>
<td>0.350</td>
<td>0.360</td>
<td>0.317</td>
<td>0.313</td>
<td>0.323</td>
<td>0.310</td>
</tr>
</tbody>
</table>

Average Closeness: 0.32%
Mean-Scores

<table>
<thead>
<tr>
<th>Year</th>
<th>T1</th>
<th>T0</th>
<th>T1</th>
<th>T2</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>0.00113</td>
<td>0.00873</td>
<td>0.00092</td>
<td>0.00024</td>
<td>0.00888</td>
<td>0.00005</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>0.00085</td>
<td>0.00766</td>
<td>0.00050</td>
<td>0.00010</td>
<td>0.00444</td>
<td>0.00002</td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>0.00073</td>
<td>0.00754</td>
<td>0.00045</td>
<td>0.00009</td>
<td>0.00325</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0.00055</td>
<td>0.00773</td>
<td>0.00036</td>
<td>0.00008</td>
<td>0.00268</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>0.00056</td>
<td>0.00766</td>
<td>0.00030</td>
<td>0.00006</td>
<td>0.00231</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>0.00045</td>
<td>0.00686</td>
<td>0.00025</td>
<td>0.00005</td>
<td>0.00179</td>
<td>0.00002</td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>0.00041</td>
<td>0.00645</td>
<td>0.00020</td>
<td>0.00005</td>
<td>0.00164</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>0.00034</td>
<td>0.00544</td>
<td>0.00018</td>
<td>0.00004</td>
<td>0.00137</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>0.00033</td>
<td>0.00478</td>
<td>0.00014</td>
<td>0.00003</td>
<td>0.00112</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>0.00028</td>
<td>0.00457</td>
<td>0.00014</td>
<td>0.00002</td>
<td>0.00094</td>
<td>0.00001</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>0.00033</td>
<td>0.00416</td>
<td>0.00012</td>
<td>0.00002</td>
<td>0.00083</td>
<td>0.00000</td>
<td></td>
</tr>
</tbody>
</table>

- Betweenness for all groups declines heavily
- T1 further improve relative superior broker position
Connectivity and Hierarchy to Tier-0 Organizations

- T1-Nodes are most direct connected to T0-Nodes and connectivity increases over time.
- T2-Nodes are weakly connected to T0-Nodes and cannot improve their situation over time.
- The 2\textsuperscript{nd} order connectivity of T1 to T0 is remarkably stable. They reach nearly half of T0-Nodes within two edges. T2 connectivity improves in this respect which is a first hint for better connectedness to T1.

The direct connection can be interpreted as degree; the indirect connection reports the average of the 2\textsuperscript{nd} order nodes with a maximum of 2 edges between origin and target node.
Connectivity and Hierarchy to Tier-1 Organizations

- T1-Nodes connect better to other T1-Nodes over time.
- Non-HEI Organizations are increasingly unimportant for T1
- Surprisingly, direct connection of T2 to T1 decreases
- The improvement of the T1 reciprocal connectedness seems to improve the 2nd order connectedness of T2

the direct connection can be interpreted as degree; the indirect connection reports the average of the 2nd order nodes with a maximum of 2 edges between origin and target node
Connectivity and Hierarchy to Tier-2 Organizations

The direct connection can be interpreted as degree; the indirect connection reports the average of the 2\textsuperscript{nd} order nodes with a maximum of 2 edges between origin and target node.

Direct connections to higher levels T0 and T1 is decreasing or at best stable, and so is the intra-tier-connectedness.

The improving 2\textsuperscript{nd} order connectedness further supports the notion of a hierarchical system with T1 brokering between global and local organizations.
Akteure der Forschungs- und Technologiepolitik in China

Hypothesen:

Aufgrund der Restrukturierung und des aktuellen Sets der Akteure bildet sich intra-regional (innerhalb Chinas) eine hierarchische Vernetzungsstruktur aus.

T1-Spitzenuniversitäten verbinden das globale mit dem lokalen Wissensnetzwerk

Die Einbindung des chinesischen in das globale Netzwerk verbessert sich im Zeitverlauf

Quelle: Hennemann/Kroll 2008:13
Networking and integration into knowledge systems both intraregional and interregional still relatively unknown.

Suitable proxy indicator?

Bibliographical information

Quelle: Hennemann/Kroll 2008:14
Degree Centrality
Describes a node’s relative or local centrality in numbers of edges from the node to others standardized by the total number of nodes in the whole network.

18 Nodes
19 Edges

\[ d = 0,29 \]
\[ d = 0,18 \]
\[ d = 0,12 \]
Closeness Centrality

Describes a node's minimal distance to all other nodes in the network. The highest closeness value in the network represents a place where time to spread information to the whole network is minimized.
Betweenness Centrality (Brandes Algorithm)

A node’s betweenness represents the node’s capability to broker between other parts of the network. It is a measure of throughput and the higher the more connections from all nodes to all nodes lead through the analyzed node.
Indirect connectivity (2nd order reachability) from one tier to another

- T1 nodes reach on average 3 T0 nodes within two steps
- 3 T0 nodes represent 37.5% of all T0 nodes