Seminar „Mustererkennung mit syntaktischen und graphenbasierten Methoden“

Thema 8
Ein Schätzverfahren für das Wachstum von Gefäßnetzwerken auf der Grundlage von Zufallsgraphen

Stefan Böcker

1. Einleitung

- Gefäßnetzwerke wachsen durch **Angiogenese** (d.h. Wachstum kleiner Blutgefäße aus einem vorgebildeten Gefäßsystem durch Sprossung oder Teilung)

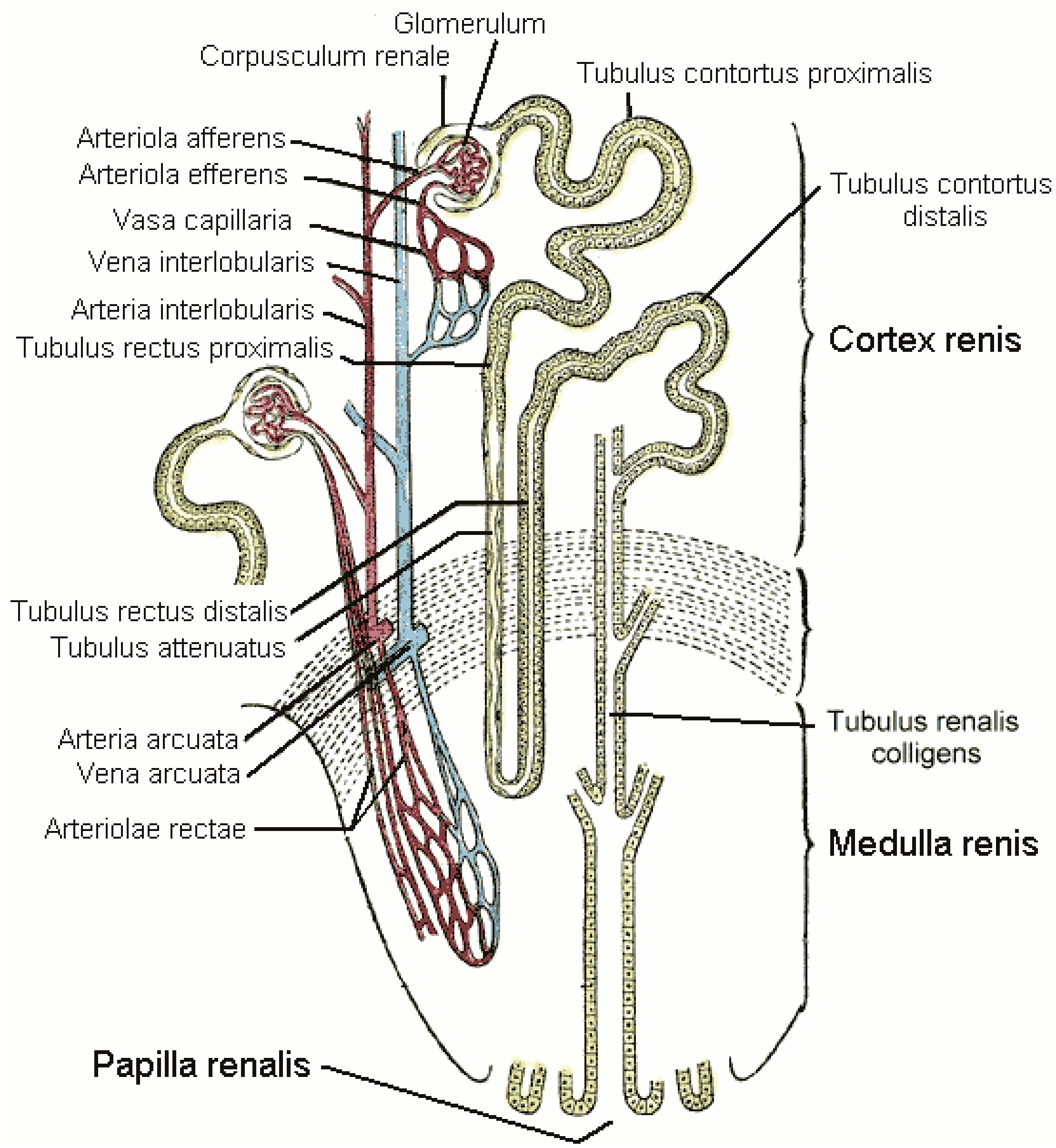
- Graphentheorie gut geeignet für Modellierung und Analyse dieser Netzwerke

- Mit Hilfe von Zufallsgraphen läßt sich Wachstum simulieren

- Als Netzwerk wird hier das **Nierenkörperchen** genutzt und untersucht

- Ziel: Schätzen der relativen Auftrittshäufigkeit von Sprossung/Teilung während des Wachstums

Niere



Quelle: http://de.wikipedia.org/wiki/Bild:Kidney_nephron.png

2. Definitionen aus Graphentheorie

2.1 Grundlagen

- Graph G besteht aus nichtleerer Knotenmenge V und Kantenmenge E, welche Paare von Knoten enthält:

$$G = (V, E)$$

- Für Gefäßnetzwerk: Kanten sind Gefäße; Knoten sind Gefäßverbindungen

- Ein *Zufallsprozess* wird beschrieben durch Menge möglicher Ergebnisse (Zustände) und den Wahrscheinlichkeiten für einen Zustandswechsel

- Ein *Schritt* entspricht einem Wechsel in einen neuen Zustand

- Ein *Zufallsgraphprozess* ist ein Zufallsprozess in welchem die Graphen Zustände sind

- Zwei Mechanismen beschreiben hierbei die Schritte:
  Teilen und *Verbinden*

- Somit läßt sich das Wachstum eines Gefäßnetzwerkes simulieren

2.2 Teilen von Gefäßen (1)

- Begonnen wird stets mit zwei Knoten, welche jeweils einem zuführenden und wegführenden Gefäßknoten entsprechen

$$E_1=\{(u, v)\}$$
$$V_1=\{u, v\}$$
$$RF_1=0$$

- In Natur existiert Sprossung und Teilung. Diese werden hier durch Verbindung und Teilung simuliert.

- Ein Zustand wird hier beschrieben als:

$$S_i=(V_i, E_i, RF_i)$$

- i endspricht der Anzahl Schritte, V ist Menge der Knoten, E ist Menge der Kanten, RF ist die relative Auftrittshäufigkeit der Teilung nach i Schritten (d.h. Anzahl Teilungen / Anzahl Schritte)

2.2 Teilen von Gefäßen (2)

- Gegeben sind also ein Zustand mit einem Graphen

$$S_i = (V_i, E_i, RF_i)$$
$$G_i = (V_i, E_i)$$

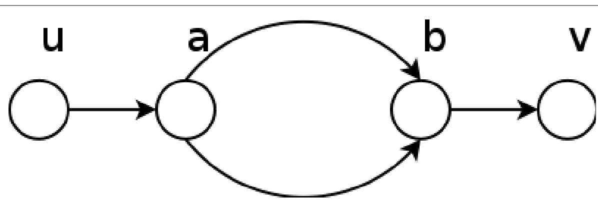
- Der Graph hat die Ordnung (Anzahl Knoten) n und die Größe (Anzahl Kanten) t

- Wähle nun eine Kante (u, v) mit Wahrscheinlichkeit 1/t und führe eine Teilung durch, um Zustand S_{i+1} mit Schritt i+1 zu erreichen

$$E_{i+1} = (E_i - (u, v)) \cup \{(u, a), (a, b), (a, b), (b, v)\}$$
$$a, b \notin V_i, (u, v) \in E_i$$

$$V_{i+1} = V_i \cup \{a, b\}$$

$$RF_{i+1} = \frac{RF_i * i + 1}{i+1}$$

S_i	S_{i+1}
	

2.3 Verbinden von Gefäßen (1)

- Die Verbindung erfolgt in zwei Teilschritten. Zunächst analog zur Teilung eine Kante wählen und Sprossung durchführen

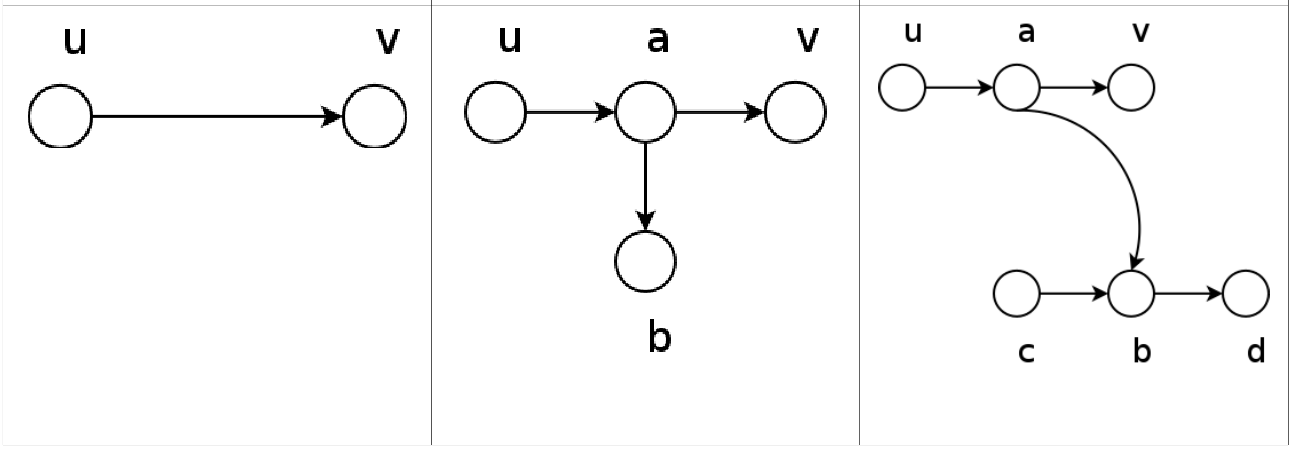
$$E'_{i+1} = (E_i - (u,v)) \cup \{(u,a),(a,b),(a,v)\}$$
$$a,b \notin V_i,\ (u,v) \in E_i$$

$$V'_{i+1} = V_i \cup \{a,b\}$$

- Von diesem Zwischenschritt mit Ordnung n+2 und Größe t+2 aus nun eine Kante (c, d) mit Wahrscheinlichkeit 1/(t+2) wählen und Knoten b mit dieser Kante verbinden

$$E_{i+1} = (E'_{i+1} - (c,d)) \cup \{(c,b),(b,d)\} \quad (c,d) \in E'_{i+1}$$
$$V_{i+1} = V'_{i+1}$$
$$RF_{i+1} = \frac{RF_i * i}{i+1}$$

S_i	S'_{i+1}	S_{i+1}



2.3 Verbinden von Gefäßen (2)

- Es ergeben sich drei wichtige Fakten:

  1. Die Anzahl der Knoten in einem generierten Zufallsgraphen nach i Schritten beträgt

  $$|V_i| = 2\mathrm{i} + 2$$

  2. Die Anzahl der Kanten in einem generierten Zufallsgraphen nach i Schritten beträgt

  $$|E_i| = 3\mathrm{i} + 1$$

  3. Ist die relative Häufigkeit der Teilung RF_i nach i Schritten, dann ist die relative Häufigkeit der Verbindung $1 - RF_i$ nach i Schritten

- Diese Fakten spielen später noch eine wichtige Rolle

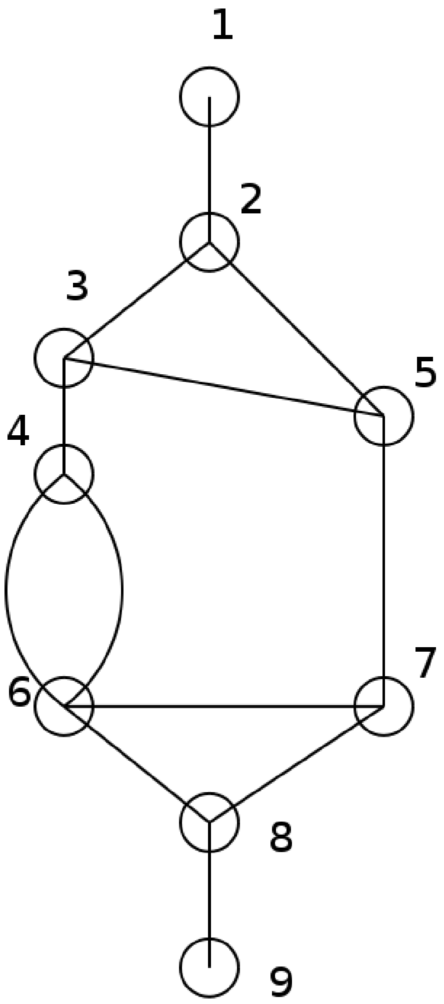
2.4 Grapheninvarianten

- Die Invariante eines Graphen beschreibt ein bestimmtes Merkmal

- formal:

$$Abbildung\ f: M \rightarrow R\ mit\ f(G) = f(G')\ falls\ f\ für\ G, G'\ definiert$$
$$und\ G, G'\ sind\ isomorph.\ M\ ist\ Teilmenge\ aller\ Graphen.$$

- wichtige Grapheninvarianten: Radius (R), Durchmesser (D) und Wurzeldistanz (rd)



	1	2	3	4	5	6	7	8	9
1	0	1	2	3	2	4	3	4	5
2	1	0	1	2	1	3	2	4	5
3	2	1	0	1	2	2	3	3	4
4	3	2	1	0	3	1	2	2	3
5	2	1	2	3	0	2	1	2	3
6	4	3	2	1	2	0	1	1	2
7	3	2	3	2	1	1	0	1	2
8	4	3	3	2	2	1	1	0	1
9	5	5	4	3	3	2	2	1	0
Exzentrizität	5	5	4	3	3	4	3	4	5

Distanzmatrix

n=9, t=12, R=min(Exz)=3,
D=max(Exz)=5, rd=d(1,9)=5

3. Schätzalgorithmus (1)

- Ziel: Schätzung des Verhältnisses von Sprossung zu Teilung während des Wachstumsprozesses eines Gefäßnetzwerks

- Input: $G_q = (V_q, E_q) \quad RF_q\, unbekannt$
- Output: RF_q

- Arbeitsweise des Algorithmus:

1. Schritt:
Generierung einer Menge von Zufallsgraphen, deren Größe der Größe von G_q entspricht.

- *Problem:* Welche Schrittanzahl wird zur Erzeugung der Zufallsgraphen gewählt?

- Zwei Möglichkeiten:

$$n_v = \frac{|V_q| - 2}{2} \qquad\qquad n_e = \frac{|E_q| - 1}{3}$$

- Beide Varianten sind ungenau, denn entweder Größe oder Ordnung der Graphen könnten sich von G_q unterscheiden

- *Lösung:*

$$n = arg\,min \sqrt{(2\mathrm{n}_i + 2 - |V_q|)^2 + (3\mathrm{n}_i + 1 - |E_q|)^2}$$
$$n_i \in \{1, \ldots, |E_q|/3\}$$

3. Schätzalgorithmus (2)

- Nun werden mittels der Schrittanzahl m Zufallsgraphen erzeugt

$$
\begin{array}{l}
\textit{for } pr = 1\% \text{ to } 100\% \quad \textit{incrementing by } 100/m\% \\
\quad S_1 = (V_1, E_1, 0)\ \textit{wobei } V_1 = \{1,2\}\ \textit{und } E_1 = \{(1,2)\} \\
\quad \textit{for } i = 2 \text{ to } n \\
\quad\quad \textit{Erweitere den Graphen mittels Teilen oder Verbinden} \\
\quad\quad\quad\quad \textit{mit der Wahrscheinlichkeit } pr \\
\quad\quad S_i = \textit{GetNextState}(S_{i-1}, Pr) \\
\quad \textit{end} \\
\quad \textit{speichere } S_n = (V_n, E_n, RF_n)\ \textit{ab} \\
\textit{end}
\end{array}
$$

2. Schritt:
Den zu $G_q = (R_q, D_q, rd_q, RF_q)$ ähnlichsten Zufallsgraphen $G_i = (R_i, D_i, rd_i, RF_i)$ finden.

$$
RF_q = arg\,min \sqrt{(R_i - R_q)^2 + (D_i - D_q)^2 + (rd_i - rd_q)^2}
$$
$$
RF_i \in \{RF_1, \ldots, RF_m\}
$$

4. Experimente / Beobachtungen

- Algorithmus basiert auf der These, dass Korrelation zwischen Grapheninvarianten und der relativen Häufigkeit der Teilung besteht

- Test mit m = 1000 und n = 80

- Beobachtung: Mit wachsender Teilungswahrscheinlichkeit, steigt auch Größe der Invarianten (Radius, Durchmesser, Wurzeldistanz)

- Dies trifft zu, wenn n > 40 und m > 100

- Experiment mit Ratten:

Studie	|V|	t	R	D	rd	n
Shea	195	322	9	15	11	96
Remuzzi	247	403	9	15	12	122
Nyengaard	256	426	9	15	11	127
Antiga	302	460	12	18	12	150
Winkler	312	466	11	16	13	155
Wahl	358	595	12	19	11	178
Urämisch	159	276	9	17	7	78
Neugeborene	24	35	4	7	5	11

- Ergebnis:

normale ausgewachsene Ratten: $RF_q = 50\%$
kranke Ratte: $RF_q > 50\%$

5. Zusammenfassung

- Ziel war das Ermitteln der relativen Häufigkeit der Gefäßteilung für ein gegebenes Blutgefäßnetzwerk

- Gefäßnetzwerke können mittels Graphen repräsentiert werden

- Mittels Zufallsgraphen lässt sich das Wachstum simulieren

- Die relative Häufigkeit der Gefäßteilung ist für die Zufallsgraphen bekannt

- Schrittanzahl zur Generierung der Zufallsgraphen wird durch Gefäßnetzwerk vorgegeben

- Generierung viele Zufallsgraphen mit gleicher Schrittanzahl

- Ermittlung der relativen Häufigkeit der Gefäßteilung

- Es besteht ein Zusammenhang zwischen Grapheninvarianten und der relativen Häufigkeit der Gefäßteilung

- **Folglich lässt sich das Verhältnis Sprossung / Teilung (bzw. Verbindung / Teilung) eines vorliegenden Gefäßnetzwerkes schätzen**

Quellenverzeichnis

Sung-Hyuk Cha / Michael L. Gargano / Louis V. Quintas / Eric M. Wahl: A vascular growth estimation algorithm using random graphs. LNCS 3434 (2005), S. 45-53.