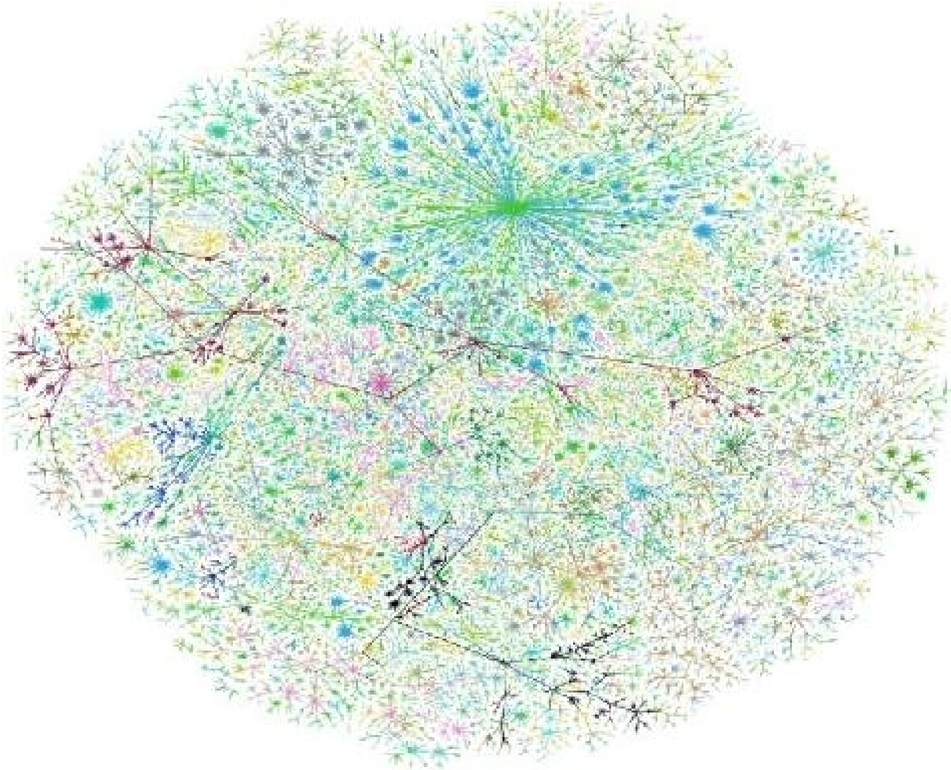
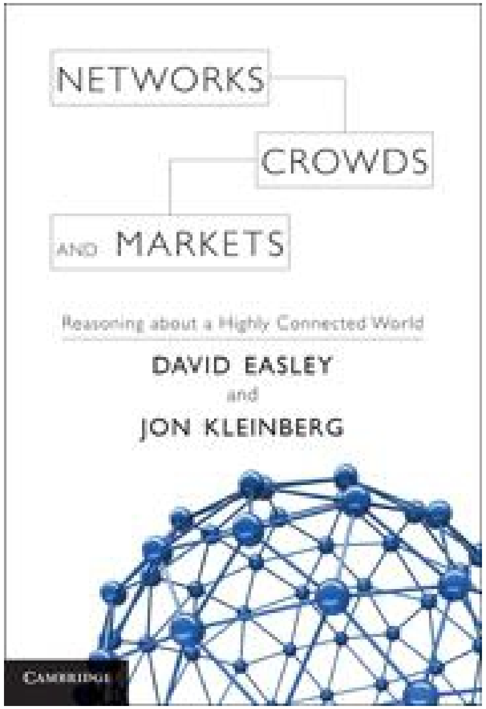
Cascading Behavior in Networks

TECHNISCHE UNIVERSITÄT DARMSTADT

Seminar aus maschinellem Lernen
WS 2010/2011





Gliederung

- Verhalten
- Ursachen für Verhaltenswahl
- Modell für Verhaltensausbreitung
- Cluster als Hindernisse
- Weak Ties
- Cascade Capacity
- Mischverhalten
- Kollektives Agieren
- Gemeinsames Wissen

Verhalten

Unterscheidbar in wesentlichen Bereichen:

- Politik
- Religion
- Nutzung von Innovationen/Technologie
- Nutzung von Kommunikationsformen(Sprache, Handynetzanbieter,....)
- ....

Ursachen für Verhaltenswahl

TECHNISCHE UNIVERSITÄT DARMSTADT

Warum entscheiden wir uns für ein bestimmtes Verhalten?

→ *Absicht zur Nutzenmaximierung*

Der (vermutete) Nutzen eines Verhaltens steht meist in Zusammenhang mit der Menge der Personen, die dieses Verhalten bereits anwenden.

Im Allgemeinen: Je mehr Menschen dieser Menge beitreten, desto eher wird man ihr selbst beitreten.

- Beispiel: Faxmaschine

Ursachen für Verhaltenswahl

→ Ursachen für Nachahmung von Verhalten unterscheidbar:

- direkter Gewinn
  - z.B. bei Verwendung kompatibler Technologien, Sprechen gleicher Sprachen,...
- informative Effekte
  - Studien: [Ryan-Gross 1943], [Coleman, Katz, Menzel 1966]

- nicht immer hängt Verhaltenswahl gleichermaßen von allen Knoten im Netzwerk ab → *globale Sichtweise* daher oft unzureichend
- oft stärker abhängig von Nachbarn im sozialen Netzwerk

  → *lokale Sichtweise* mit Einbezug der tieferen Netzwerkstruktur
  - z.B. von Interesse für Viral Marketing

Modell für Verhaltensausbreitung

Einführen eines Modells für Modellierung einer Verhaltensausbreitung mit lokaler Sichtweise:

- Gegeben sei soziales Netzwerk
- 2 mögliche Verhaltensarten für jeden Knoten: **A** oder **B**
- Prozess läuft schrittweise ab:
  - *0. Schritt:* alle Knoten haben **B**
  - *1. Schritt:* ein paar Knoten nehmen spontan **A** an(*Early Adopters*)
  - *ab 2. Schritt:* jeder Knoten(**außer *Early Adopters***) wählt Verhalten mit <u>größtem Payoff</u> (bei gleichem Payoff, wird **A** gewählt)

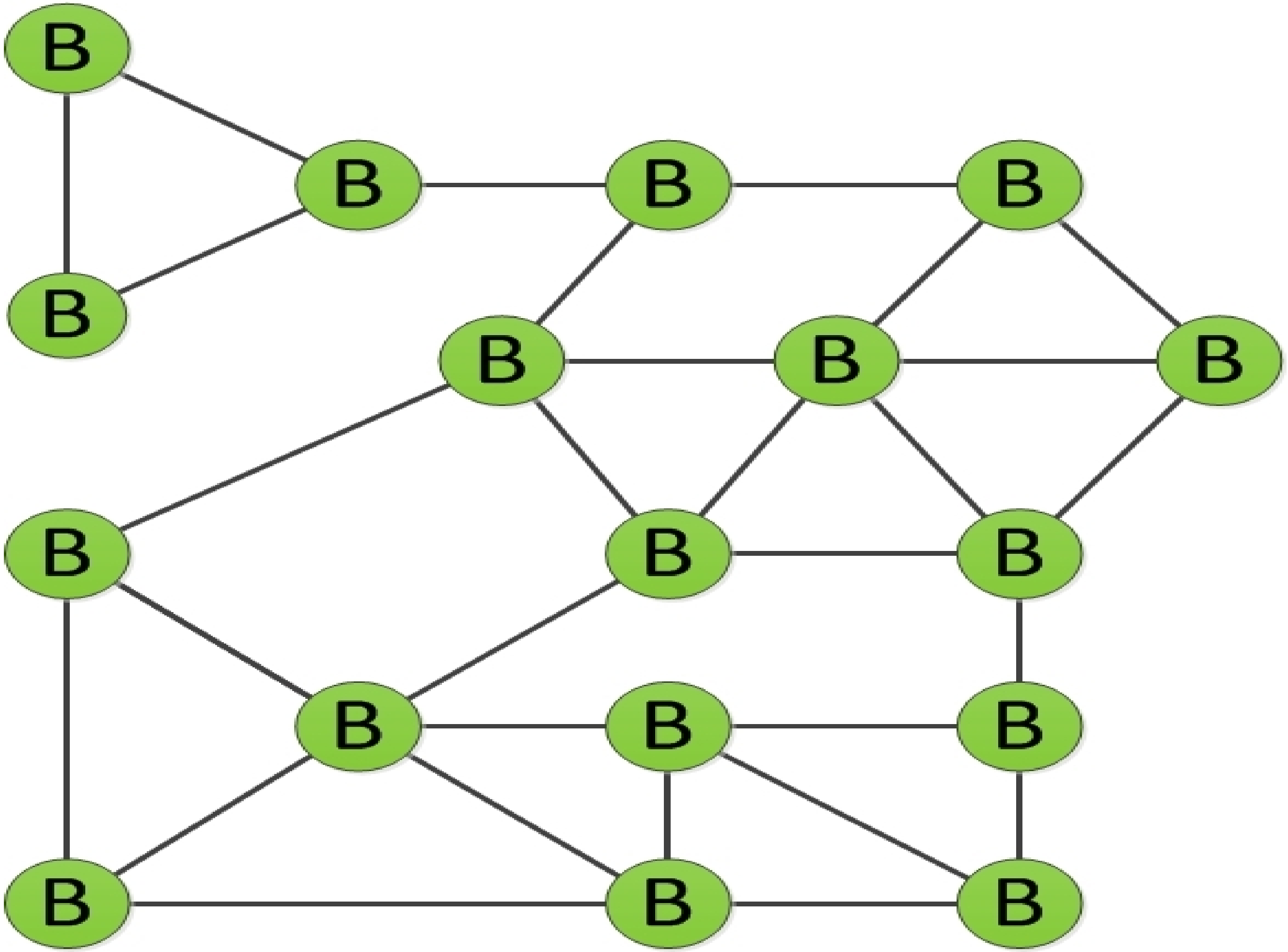
Modell: Payoffs und Threshold

- ein Knoten v erhält für jeden Nachbarn u_i , abhängig vom Verhalten beider, Payoff po_i gemäß Payoff-Matrix:

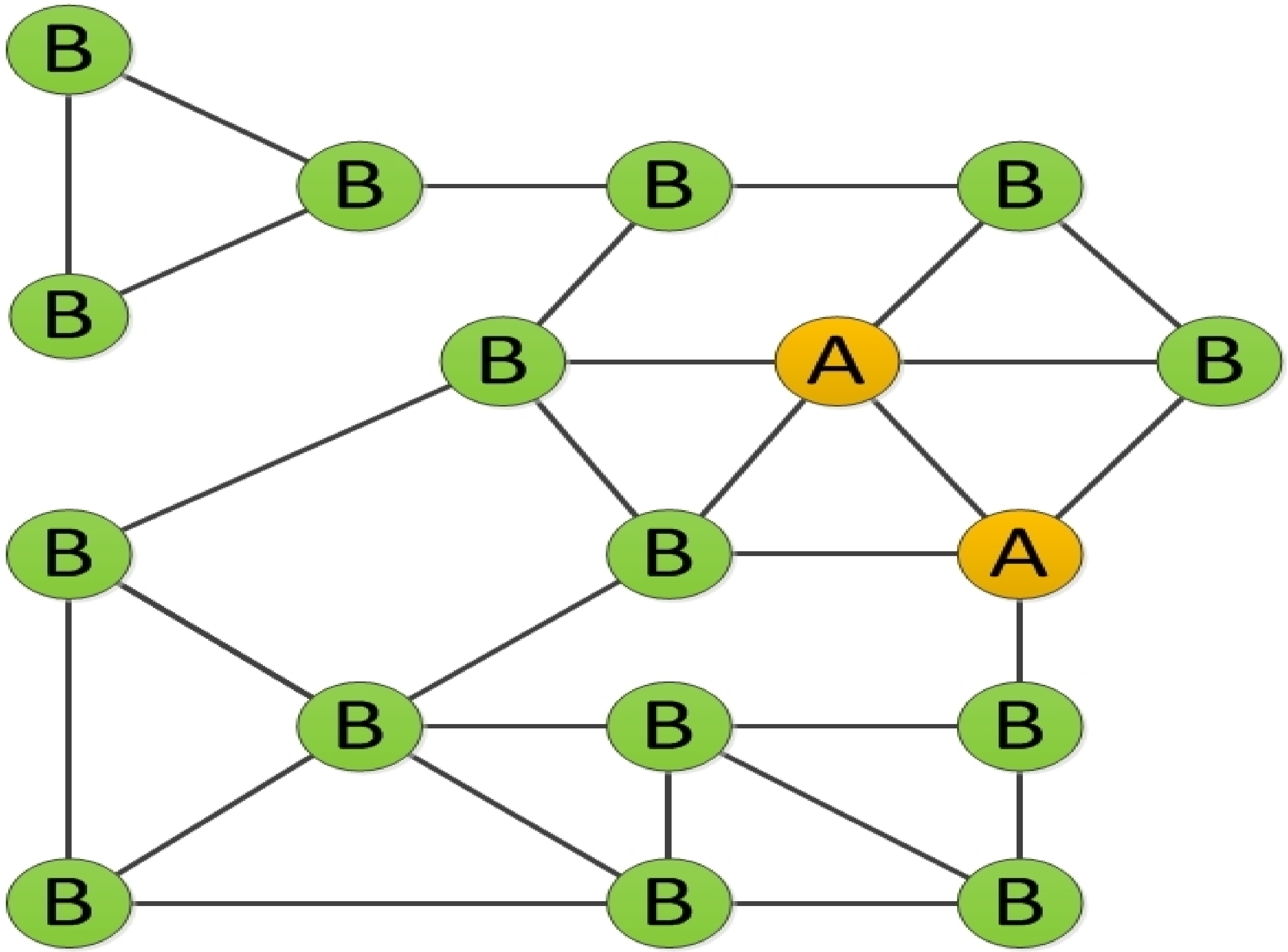
PAYOFF MATRIX	**A**	**B**
A	a,a	0,0
B	0,0	b,b

- *d:* Anzahl der Nachbarn von Knoten *v*
- → *p:* <u>Anteil</u> der Nachbarn von v mit Verhalten **A**
- v nimmt A an $\Leftrightarrow \; a \cdot d \cdot p \geq b \cdot d \cdot (1 - p) \; \Leftrightarrow \; p \geq \frac{b}{a+b} = q$
- → *q ist* Threshold für Annahme von Verhalten A
- Beispiel: *a*=3, *b=2* → $q = \frac{2}{5}$

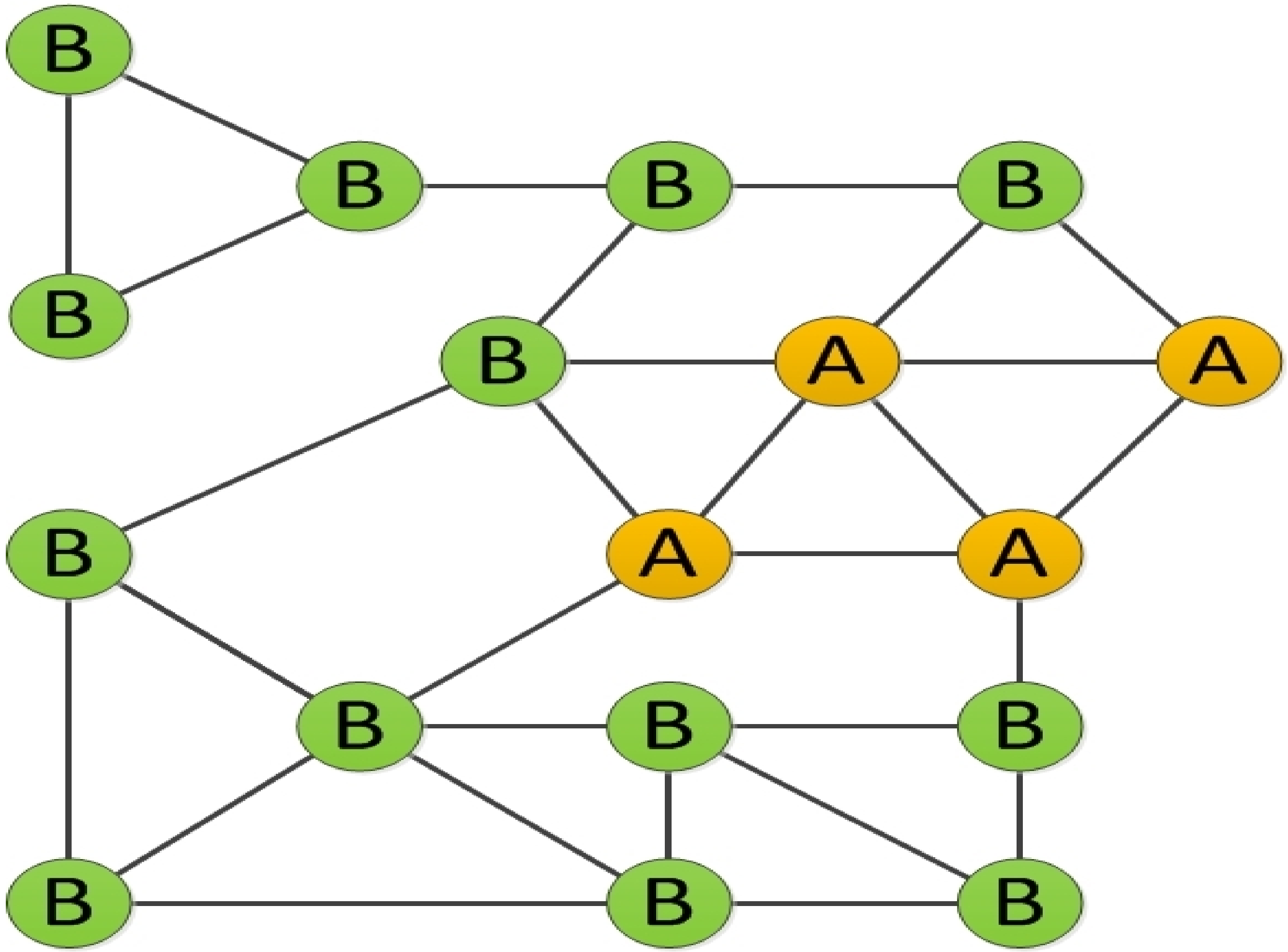
Beispiel für $q = \frac{2}{5}$ - Schritt k=0



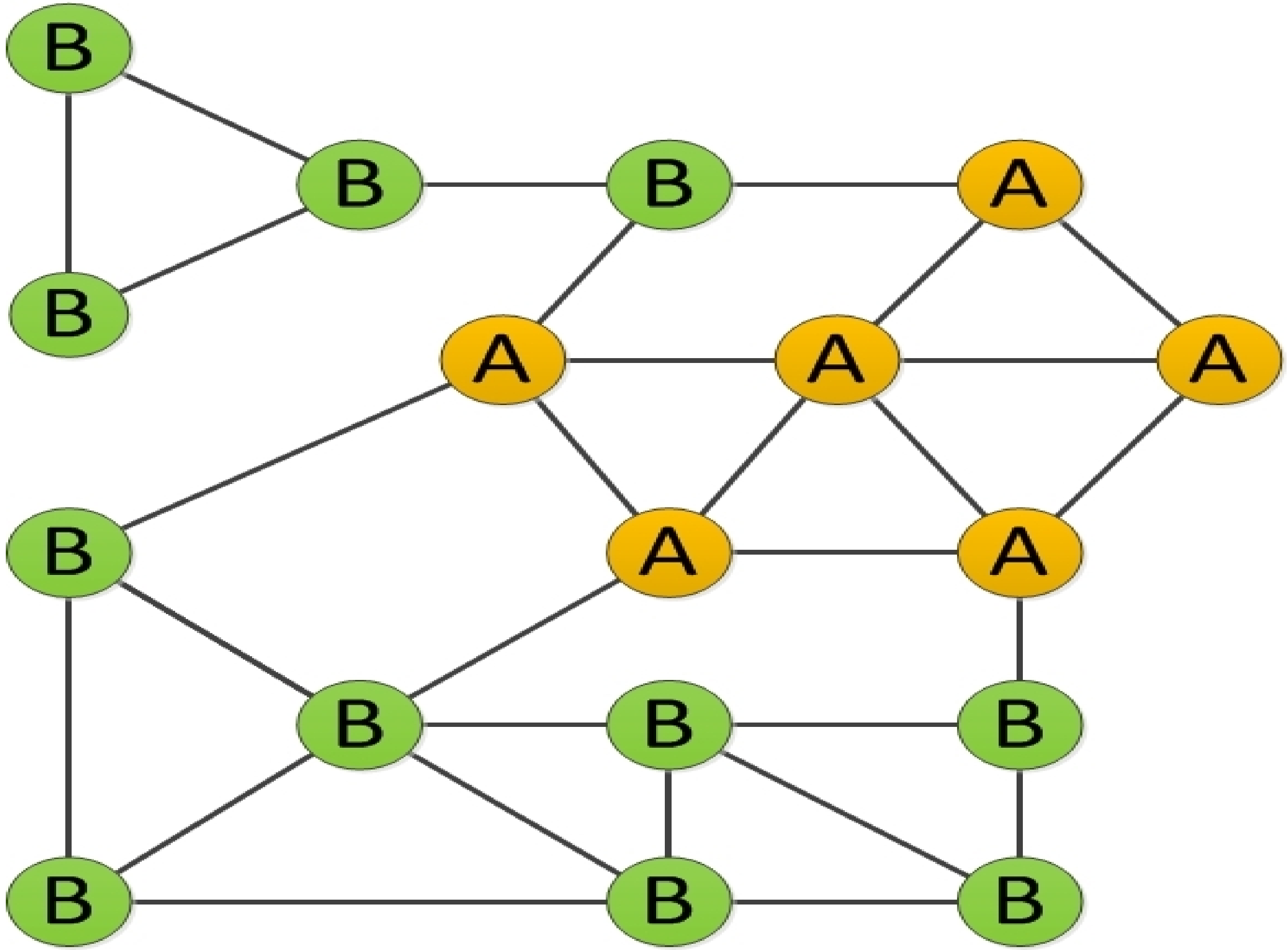
Beispiel für $q = \frac{2}{5}$ - Schritt k=1



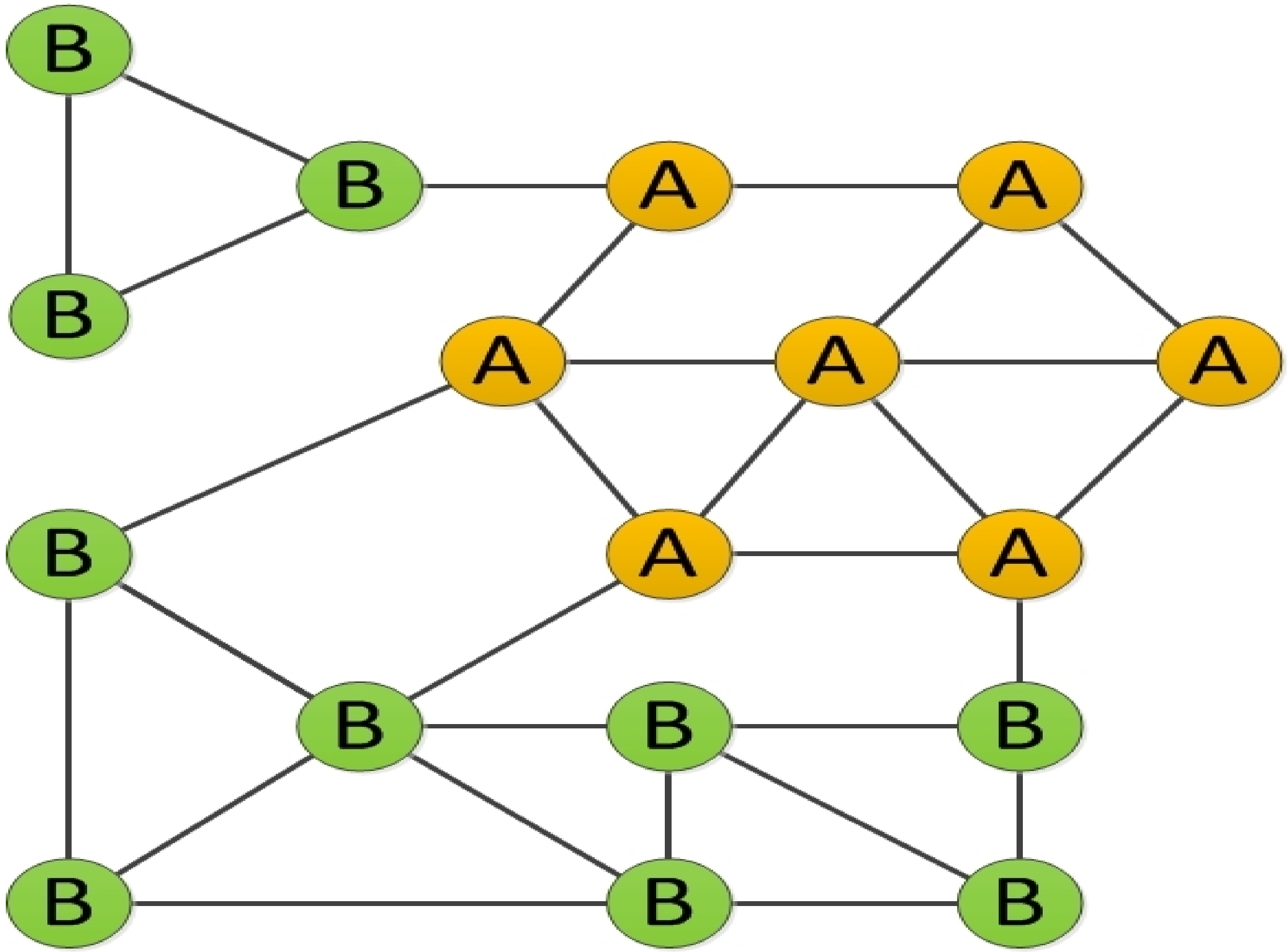
Beispiel für $q = \frac{2}{5}$ - Schritt k=2



Beispiel für $q = \frac{2}{5}$ - Schritt k=3



Beispiel für $q = \frac{2}{5}$ - Schritt k=4



Modell: Verhaltensausbreitung

Ein Knoten, der einmal Verhalten A angenommen hat, wird niemals mehr Verhalten B annehmen.

Beweis:

- Für *Early Adopters* gilt dies nach Voraussetzung
- Für die restlichen Knoten:
  - wenn in Schritt k kein Knoten von A nach B wechselt, ist dies in Schritt k+1 ebenfalls so, da *p* für jeden Knoten entweder gleich bleibt oder steigt
  - und somit gilt dies auch für alle Schritte ≥ k
  - ist der Fall ab Schritt k=1

→ *Verhalten A breitet sich ab Schritt 1 stetig aus, bis sich ein stabiler Endzustand einstellt.*

Modell: Endzustand

→ Zustand, ab dem keine Veränderung mehr stattfindet

Kann verschieden aussehen:

- 1.) A hat B vollständig verdrängt(vollständige Kaskade)
- 2.) A und B koexistieren im Netzwerk

Welcher Fall eintritt, hängt ab von:

- Netzwerkstruktur(→Cluster)
- Schwellwert q bzw. den Payoffs a und b
- Early Adopters

Cluster als Hindernisse

TECHNISCHE UNIVERSITÄT DARMSTADT

Def.: Ein Cluster der Dichte d ist eine Menge von Knoten, sodass für jeden dieser Knoten gilt, dass **mindestens** der Anteil d seiner Nachbarn in dieser Menge enthalten ist.

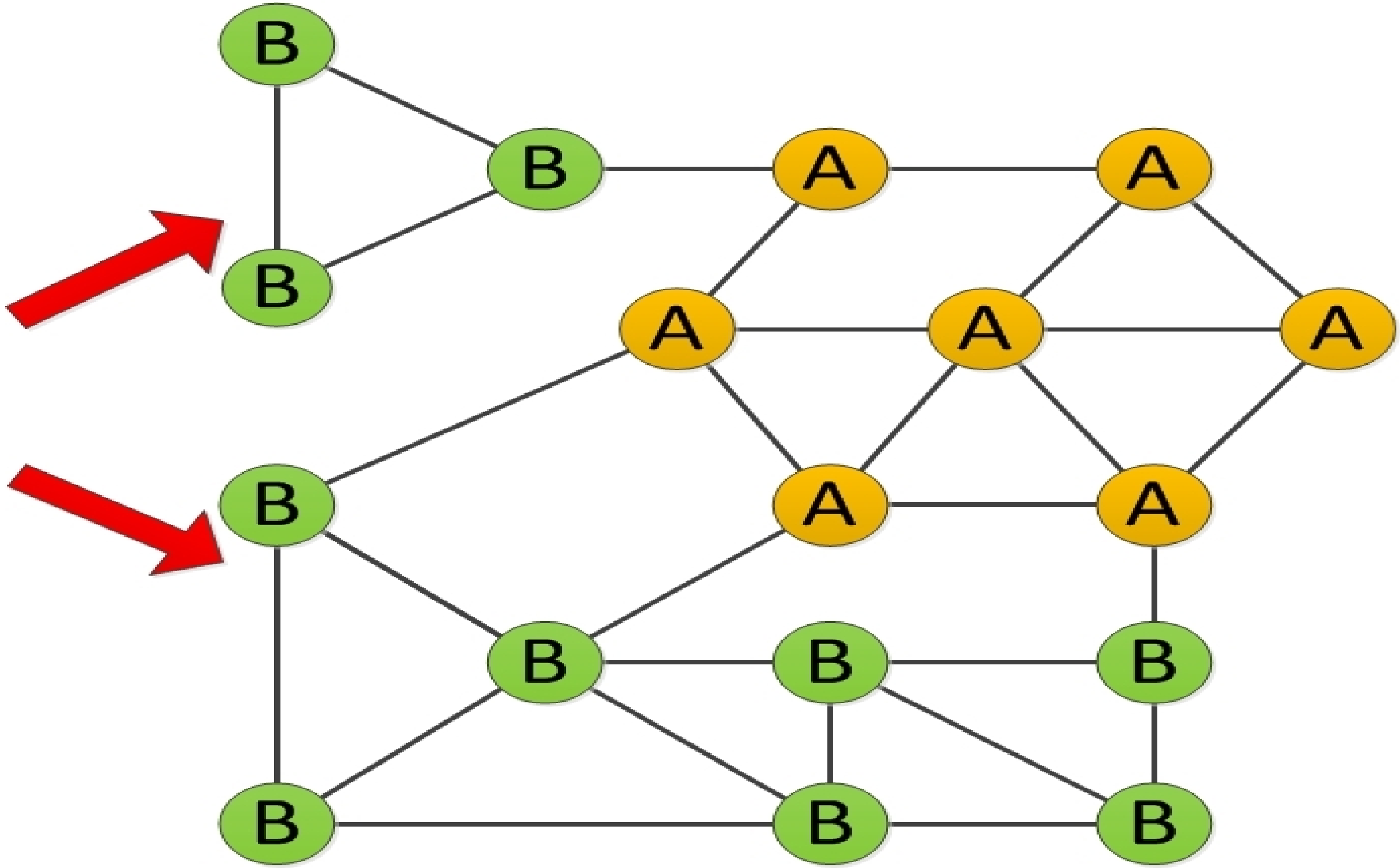
→ *Ein Cluster der Dichte d > 1 - q mit Verhalten B wird einem sich im Netzwerk ausbreitendem Verhalten A widerstehen(keiner dieser Knoten wird A annehmen).*

Beweis:

- Für ersten Knoten *v* aus Cluster, der **A** annimmt, gilt dass
  $1 - p \geq d \wedge d > 1 - q \Rightarrow p < q \Leftrightarrow$ *v* kann nicht **A** angenommen haben

Beispiel: Endzustand mit $q = \frac{2}{5}$

$d = \frac{2}{3}$



Cluster als Hindernisse

→ Umkehrung gilt auch:

Wenn die Ausbreitung eines Verhaltens A in einem Netzwerk zum Erliegen kommt, dann ist die Ursache hierfür ein Cluster mit dem Verhalten B und der Dichte $d > 1 - q$.

<u>Beweis:</u>

- verbleibende Knoten mit Verhalten B im Endzustand bilden offensichtlich einen Cluster mit der Dichte $d < 1 - q$

Cluster als Hindernisse

TECHNISCHE UNIVERSITÄT DARMSTADT

Wie „bekehrt“ man einen Cluster, wenn $d > 1 - q$?

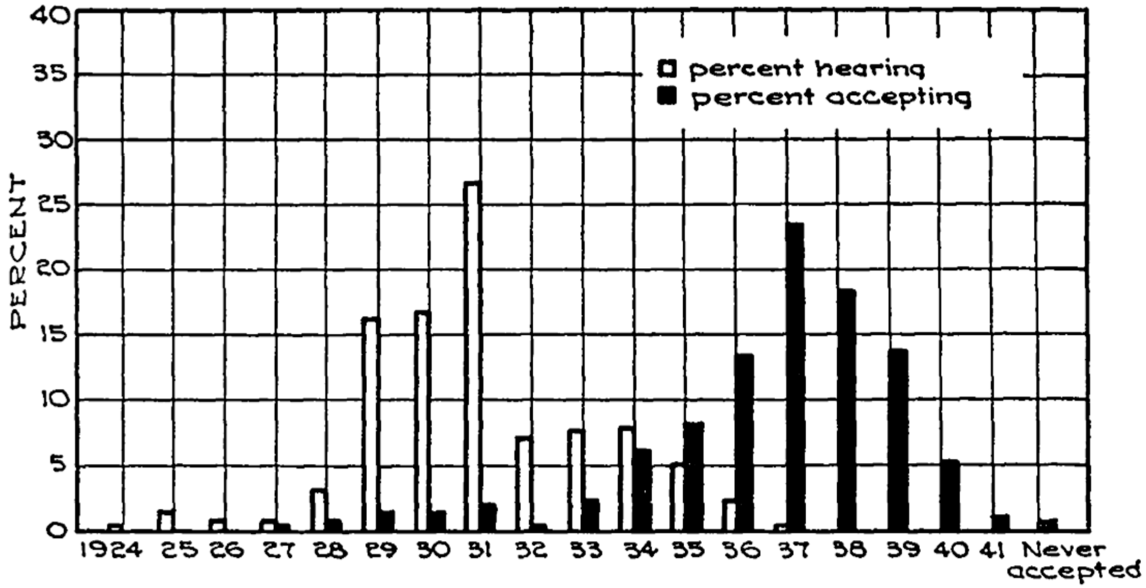
Möglichkeiten:

- Von Innen: Erzeugen von Early Adopters im Cluster(Überzeugung,....)
- Von Außen: Erhöhen des Payoffs a (z.B. Verbesserung einer Technologie), sodass *q* geringer wird

Weak Ties

→ Schwache Verbindungen zu Clustern: *Weak Ties*

- Über Weak Ties findet eher Ausbreitung von Informationen als Verhaltensweisen statt



Verbreitung von Information bzw. Verwendung des Hybrid-Samenkorns in Ryan-Gross-Studie

Cascade Capacity

Def.: *Cascade Capacity* ist größter Wert für *q* bzgl. eines Netzwerkes, für das eine kleine Menge von Early Adopters eine vollständige Kaskade darin auslösen kann.

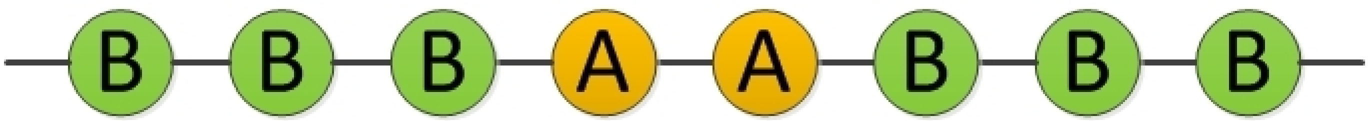
Modellannahmen:

- Netzwerkgröße: unendlich (für „sehr groß“)
- Anzahl an Early Adopters: endlich(für „klein“)

Vollständige Kaskade ⇔ jeder Knoten ändert Verhalten in endlicher Zeit

Beispiel: Unendliche Kette

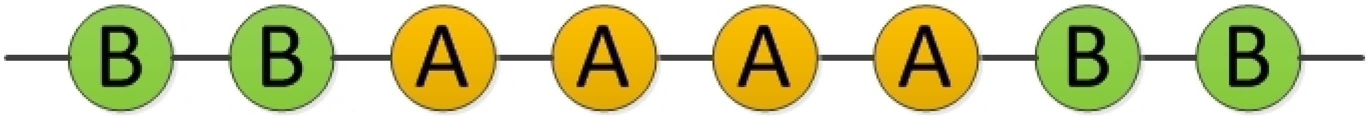
→ *Cascade Capacity* $= \frac{1}{2}$



Vollständige Kaskade nur, wenn q $\leq \frac{1}{2}$

Beispiel: Unendliche Kette

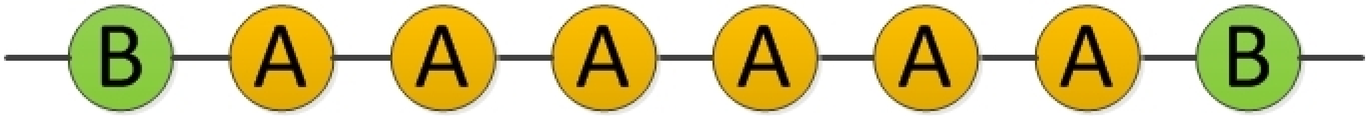
→ *Cascade Capacity* $= \frac{1}{2}$



Vollständige Kaskade nur, wenn $q \leq \frac{1}{2}$

Beispiel: Unendliche Kette

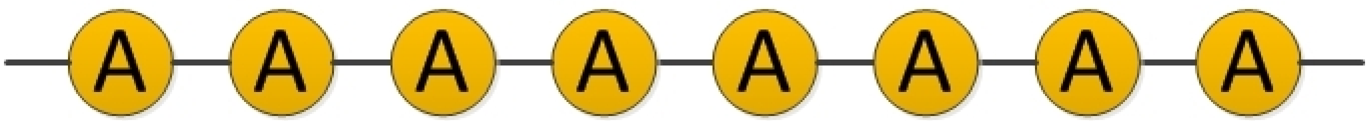
→ *Cascade Capacity* $= \frac{1}{2}$



Vollständige Kaskade nur, wenn $q \leq \frac{1}{2}$

Beispiel: Unendliche Kette

→ *Cascade Capacity* $= \frac{1}{2}$



Vollständige Kaskade nur, wenn $q \leq \frac{1}{2}$

Cascade Capacity

→ Je geringer *Cascade Capacity* ist, desto wahrscheinlicher ist eine weite bzw. vollständige Ausbreitung eines Verhaltens

Kein Netzwerk hat eine Cascade Capacity $> \frac{1}{2}$

- schlechtere Innovation wird sich gegen eine bessere weit verbreitete niemals durchsetzen

Beweis:

- Menge von **A**-**B**-Kanten ist endlich, da **A**-Knoten sowie deren Nachbarn endlich sind
- Annahme: $q > \frac{1}{2}$:
  - wenn ein **B**-Knoten **A** annimmt, verringert sich diese Menge
  - Anzahl **A**-**B**-Kanten nimmt stetig ab → Ausbreitung muss irgendwann (in endlicher Zeit) stoppen

Mischverhalten

Beispiel: Mehrsprachigkeit

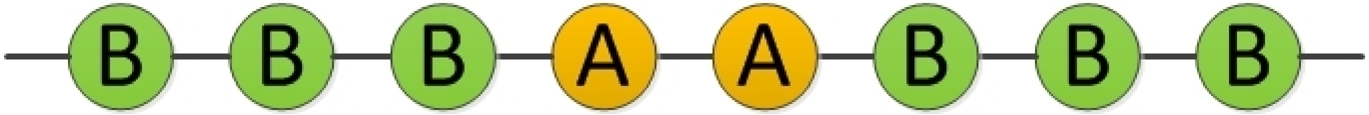
→ Einführen von Verhalten **AB** als Kombination aus **A** und **B**

- Payoff:

PAYOFF MATRIX	**A**	**B**	**AB**
A	a,a	0,0	a,a
B	0,0	b,b	b,b
AB	a,a	b,b	max(a,b), max(a,b)

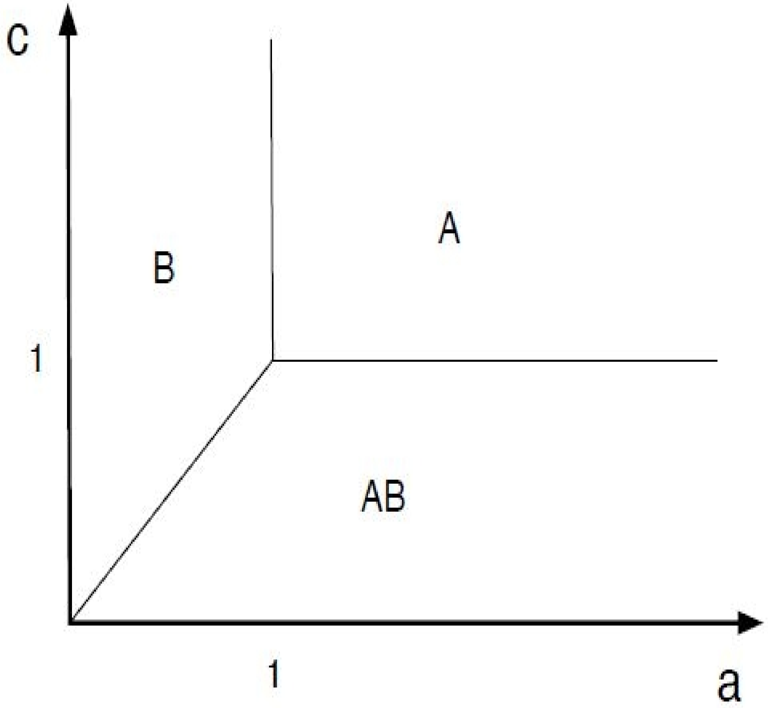
- Kosten für Anwendung von **AB**: *c*

Beispiel: Unendliche Kette

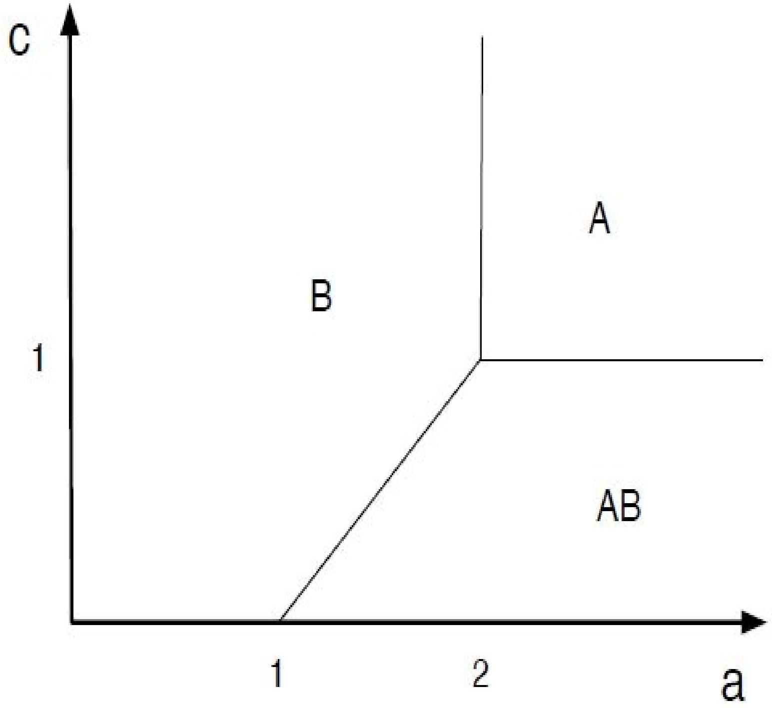


Cascade Capacity für Unendliche Kette

→ Normalisierung der Payoffs, sodass b = 1

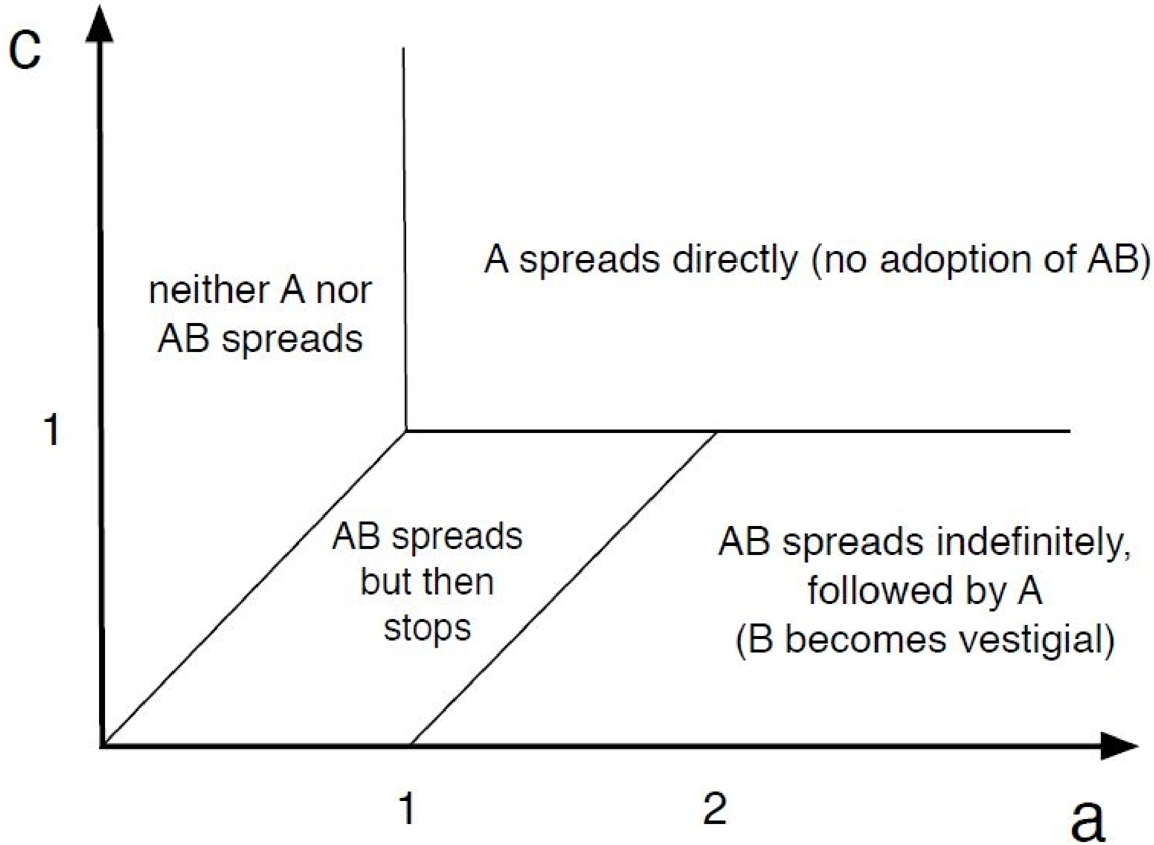


Angenommenes Verhalten, wenn Knoten zwischen **A** und **B** liegt



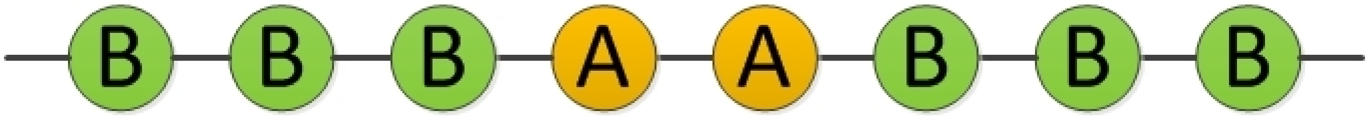
Angenommenes Verhalten, wenn Knoten zwischen **AB** und **B** liegt

Cascade Capacity mit AB



Beispiel: Unendliche Kette

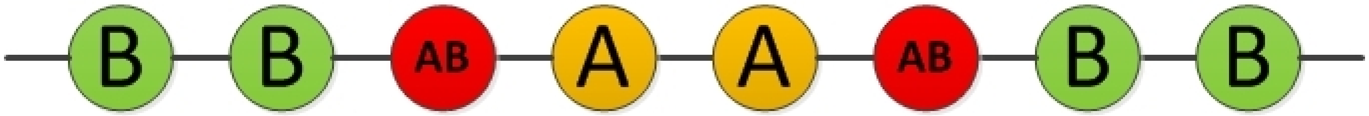
→ a = 1, b = 1, c = 0.5



Endzustand!
Pufferzone **AB** zwischen **A** und **B**.

Beispiel: Unendliche Kette

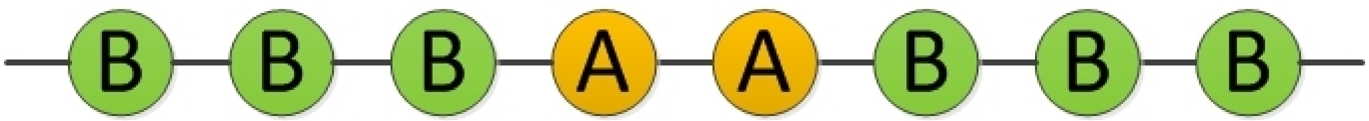
→ a = 1, b = 1, c = 0.5



Endzustand!
Pufferzone **AB** zwischen **A** und **B**.

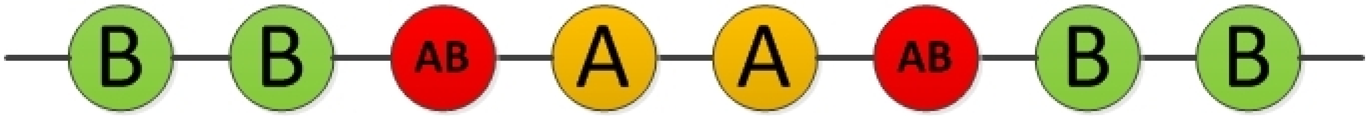
Neues Beispiel: Unendliche Kette

→ $a = 4$, $b = 1$, $c = 0.5$



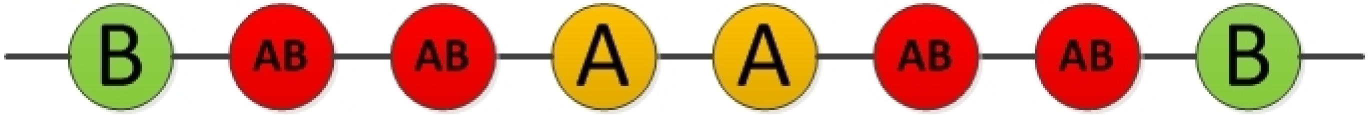
Neues Beispiel: Unendliche Kette

→ $a = 4, b = 1, c = 0.5$



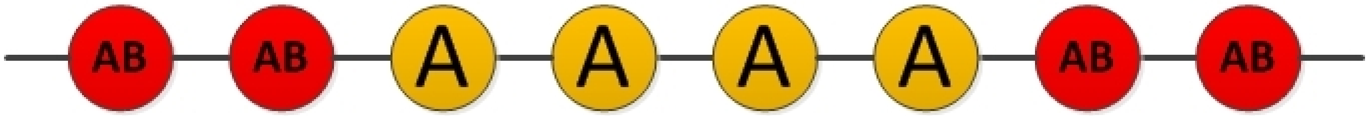
Neues Beispiel: Unendliche Kette

→ $a = 4$, $b = 1$, $c = 0.5$



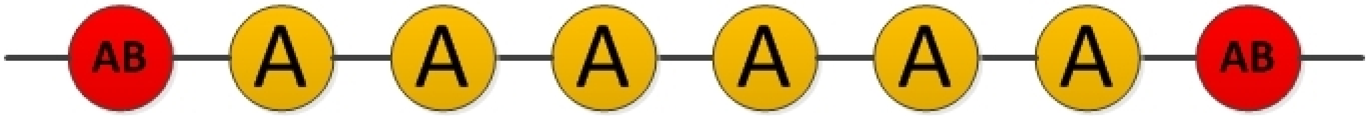
Neues Beispiel: Unendliche Kette

→ $a = 4$, $b = 1$, $c = 0.5$



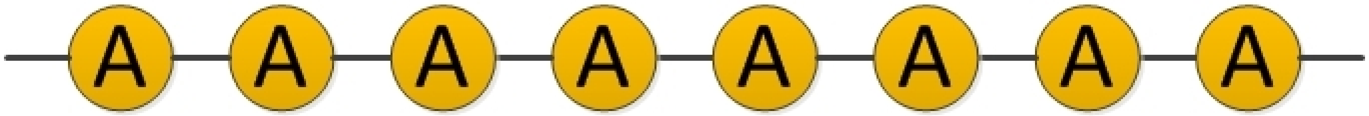
Neues Beispiel: Unendliche Kette

→ a = 4, b = 1, c = 0.5

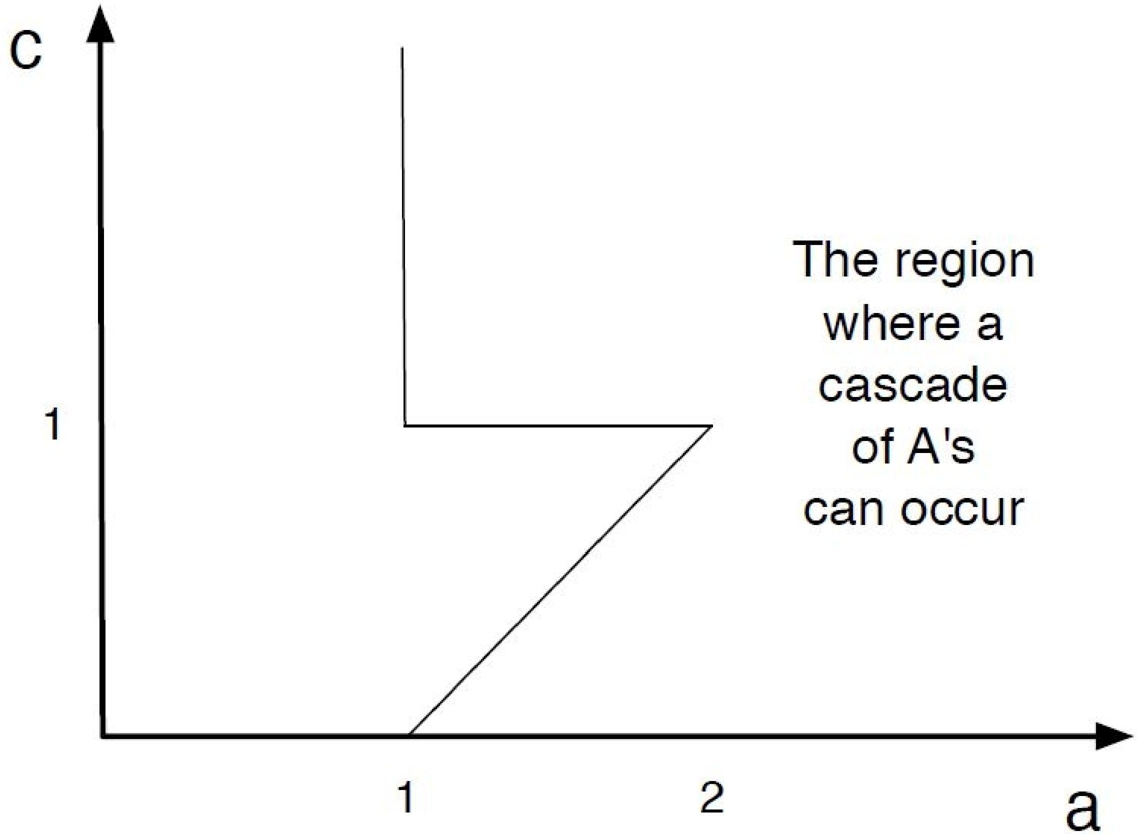


Neues Beispiel: Unendliche Kette

→ $a = 4, b = 1, c = 0.5$



Cascade Capacity mit AB



Interpretation

→ ist Payoff a nicht groß genug und sind Kosten c weder zu klein noch zu groß, wird sich eine Pufferzone in Form von AB bilden.

Manchmal kann das Eingehen eines Kompromisses bzw. das Ermöglichen von Kompatibilität mit einem anderen besseren Verhalten das Überleben eines Verhaltens ermöglichen.

Kollektives Agieren und Wissen

TECHNISCHE UNIVERSITÄT DARMSTADT

In bestimmten Fällen ist die Teilnahme an einer Aktion nur dann positiv(Payoff > 0), wenn genügend Leute daran teilnehmen, ansonsten ist sie negativ(Payoff < 0).

→ Wissen über Verhalten anderer Menschen muss vorhanden sein, d.h. Wissen muss sich ausbreiten können

- Beispiel Faxmaschine: Wissen vorhanden

Kollektives Agieren und Wissen

Problem: Informationsflüsse sind blockiert oder werden verfälscht

→ führt zu *Pluralistischer Ignoranz*

Beispiele:

- repressive Systeme
  - Mehrheit ist gegen die Regierung, niemand weiß dies jedoch
    → Regierung wird überleben
- Umfragen in USA um 1970 zu Rassentrennung
  - Nur eine Minderheit war für Rassentrennung, über 50% glaubten jedoch, es wäre die Mehrheit

Modell: Kollektives Agieren

TECHNISCHE UNIVERSITÄT DARMSTADT

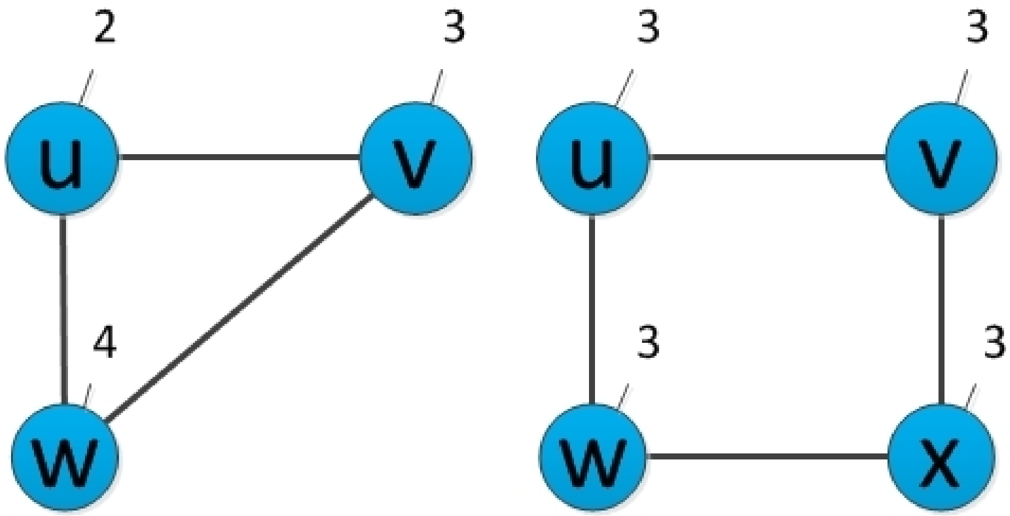
Jede Person im Netzwerk nimmt nur an einer Aktion teil, wenn sie weiß(besser: glaubt), dass eine bestimmte Mindestzahl(Threshold) an der Aktion teilnimmt.

Modellannahmen:

Jede Person kennt:

- Thresholds ihrer Nachbarn(die restlichen nicht)
- Struktur des Netzwerkes

Modell: Beispiel



Keine Aktion! **Keine Aktion!** **Aktion möglich!**

Gemeinsames Wissen

→ *Wissen, von dem jeder weiß, dass es jeder weiß.*

- Gewissermaßen Gegenteil von *pluralistischer Ignoranz*
- Wird vor Allem geschaffen durch Mainstream-Medien,...
- Bedeutung für:
  - politische Ansichten
  - Verkauf von Produkten, die nur nützlich sind, wenn viele dieses Produkt kaufen(z.B. Faxmaschine, Mobil-Telefone,....)
  - ....

Beispiele: Gemeinsames Wissen

TECHNISCHE UNIVERSITÄT DARMSTADT

- z.B. „Apple Macintosh“-Werbung bei *Super Bowl* 1984

  → galt noch Jahre danach als *“Greatest Television Commercial of All Time”*

- Religion im Irak nach US-Invasion 2003:
  - Shiiten erfolgreicher bei Erreichung von nationalen Zielen als Sunniten, da Verbreitung von Informationen mittels Moschee-Netzwerken besser zur Schaffung *gemeinsamen Wissens* geeignet war(bzgl. Wahlstrategien usw.)

Danke für die Aufmerksamkeit!

Fragen?

Quellen

- David Easly, Jon Kleinberg: Networks, Crowds and Markets: Reasoning about a Highly Connected World, 2010, Chapter 19: Cascading Behavior ind Networks
- http://www.cise.ufl.edu/AppliedOptimization/slides/leskovec_jure_01.pdf
- http://bilder.bild.de/BILD/regional/stuttgart/dpa/2010/09/05/stuttgart-21--demonstration,templateId=renderScaled,property=Bild,height=349.jpg
- http://www.uni-graz.at/yvonne.schmidt/varrel1.gif
- http://www.teltarif.de/arch/2010/kw21/editorial-kommentar-tablets-arm-netbooks-computex-1m.jpg

Zusätzliche Erweiterungen des Grundmodells

→ Einführen heterogener Thresholds q_v:

- Payoff-Matrix für jedes Paar (v,w) von Knoten:

PAYOFF MATRIX	A	B
A	a_v,a_w	0,0
B	0,0	b_v,b_w

→ führt für Knoten v zu $q_v = \dfrac{b_v}{a_v+b_v}$

Heterogene Thresholds

- Existenz leicht und schwer beeinflussbarer Menschen wird Rechnung getragen
- Nicht nur Macht an sich ist entscheidend(zentrale Position im Netzwerk), sondern auch Zugang zu leicht beinflussbaren Menschen
- Viral Marketing: Erzeugen von Early Adopters als Trade-Off zwischen ihrer zentralen Position und Zugang zu leicht beeinflussbaren Menschen

Verallgemeinerter Payoff

- Hinzufügen eines einmaligen Payoffs für jedes Verhalten
- Payoff

PAYOFF MATRIX	A	B
A	a_v,a_w	0,0
B	0,0	b_v,b_w

- **A** bringt einmalig p_A und **B** p_B
- *nehme **A** an* $\Leftrightarrow p \cdot d \cdot a_v + p_A \geq (1-p) \cdot d \cdot b_v + p_B \Leftrightarrow q_v = \dfrac{b_v + \dfrac{p_B - p_A}{d}}{a_v + b_v}$