

# Seminari LabSIMQ

*Un esempio di applicazione di un metodo di analisi  
ispirato alla complessità*

Ires, 17 dicembre 2009, ore 10 sala Rosa

Micro stime da macro indagini: un'applicazione  
degli algoritmi genetici ai tassi di adozione della  
banda larga da parte delle imprese piemontesi

*Riccardo Boero*

# Macro dati e micro stime

- Problema diffuso
  - Scienze regionali
  - Policies
- Necessità di colmare un gap informativo tramite una stima, utilizzando tutte le eterogenee fonti informative a disposizione

# Outline

- Un problema
- Un metodo
- L'implementazione
- I risultati

# Indagine oss. ICT (2005)

# addetti	adozione
10 – 49	71.9%
50 – 249	78.3%
Oltre 250	100.0%

Totale Piemonte: 75.3%

settore	adozione
NACE D	71.5%
NACE G	76.9%
NACE H	81.8%
NACE I	52.8%
NACE K	94.6%
<b>Altri</b>	<b>74.7%</b>

# Diffusione (fine I trim)

	2002	2003	2004	2005	2006	2007	2008
comuni connessi	54	106	159	288	374	738	935

# Popolazione imprese

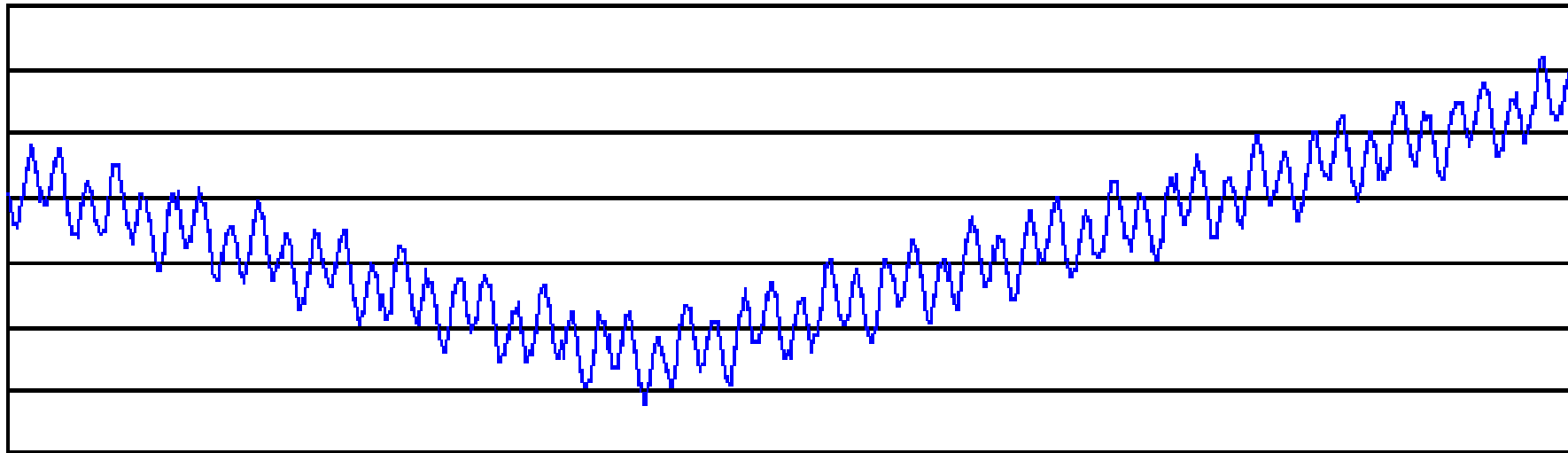
- Comune 1236:

<b>settore</b>	<b>10-49</b>	<b>50-249</b>	<b>oltre 250</b>
NACE D	9	7	0
NACE G	67	23	1
NACE H	23	33	2
NACE I	11	0	0
NACE K	0	2	1
Altri	179	56	7

# Il problema dal punto di vista matematico

- Trovare i valori delle imprese connesse per ciascun comune, classe dimensionale e settoriale.
- Si conoscono i valori totali regionali per settore e classe dimensionale, e i valori massimi raggiungibili tramite la popolazione delle imprese e i comuni raggiunti dal servizio.
- È un'ottimizzazione vincolata, ma...

# Lo spazio delle soluzioni



<http://www.obitko.com/tutorials/genetic-algorithms/search-space.php>



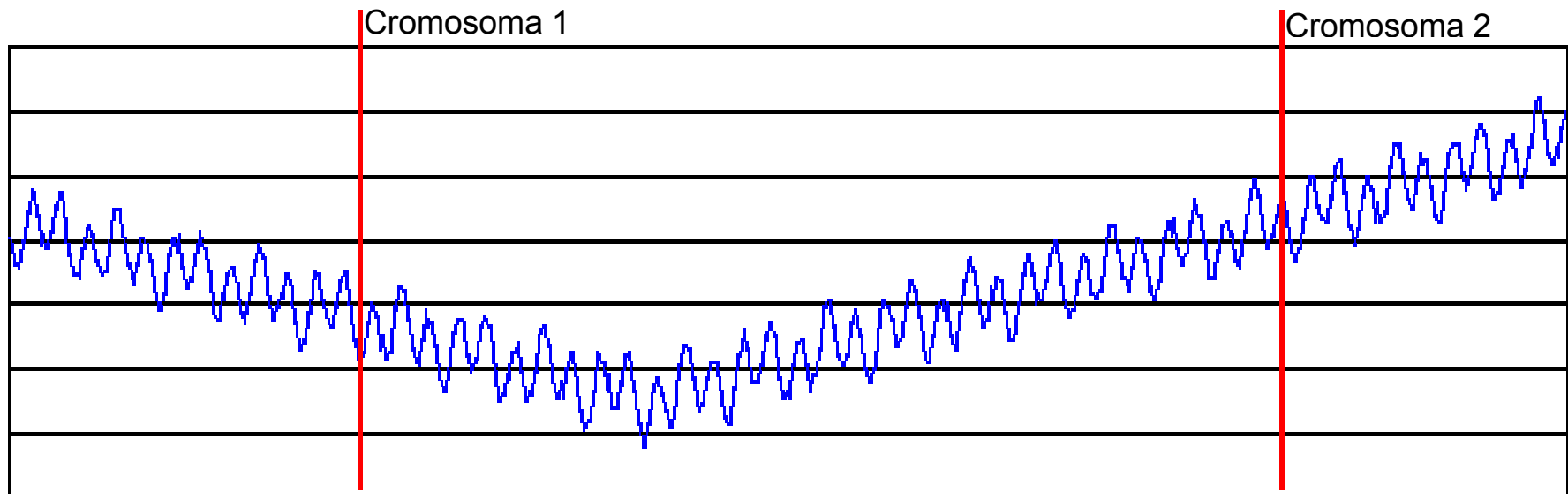
# Algoritmi genetici

- È un'euristica
- Si parla di generazioni costituite da soluzioni
- Ad esempio la prima generazione, a caso (esempio)
  - Cromosoma 1: 1101100100110110
  - Cromosoma 2: 1101111000011110

# Processo di ottimizzazione

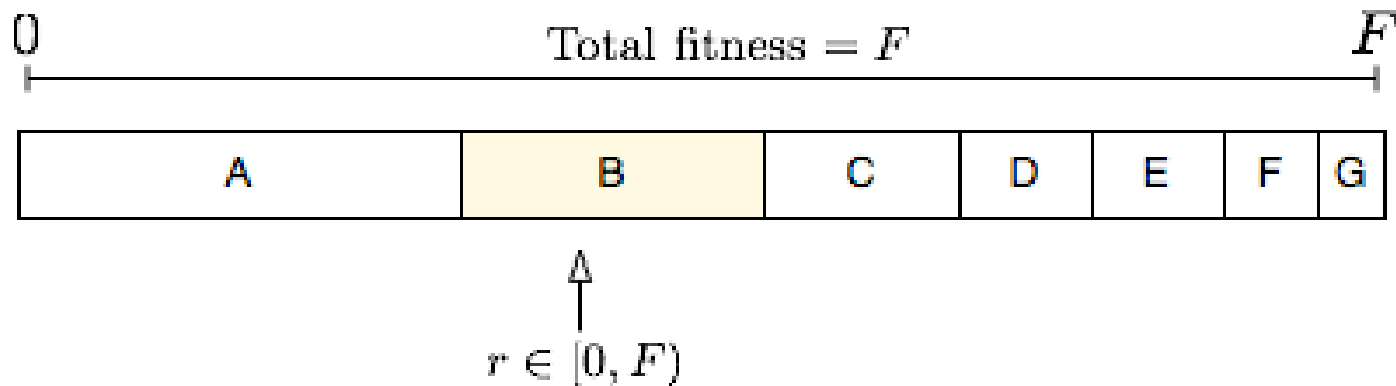
- Passaggio tramite alcuni processi a generazioni successive.
- 1: Valutazione della fitness di ciascuna soluzione nella generazione attuale
- 2: Selezione delle soluzioni più performanti e applicazione di:
  - Riproduzione
  - Crossover
  - Mutazione
- 3: Nuova generazione

# Valutazione



# Selezione

- Meccanismo a roulette wheel



[http://en.wikipedia.org/wiki/File:Fitness\\_proportionate\\_selection\\_example.png](http://en.wikipedia.org/wiki/File:Fitness_proportionate_selection_example.png)

# Crossover

- Cromosoma 1: 1101100100110110
- Cromosoma 2: 1101111000011110
- Cromosoma 1: 11011 | 00100110110
- Cromosoma 2: 11011 | 11000011110
- Figlio 1: 11011 | 11000011110
- Figlio 2: 11011 | 00100110110

# Mutazione

- Originale 1: 110**1**111000**0**11110
- Originale 2: 110110**0**100110110
- Mutato 1: 110**0**111000**1**11110
- Mutato 2: 110110**1**100110110

# Implementazione

- Il codice di ciascun cromosoma rappresenta il numero di imprese connesse in ciascun comune, per ciascun settore e classe dimensionale, per tutti i comuni
- I valori dei geni sono vincolati a essere compresi nei valori massimi (imprese presenti oppure 0 se comuni non connessi)

# Funzioni di fitness

- Size fitness: l'adozione per dimensione a livello regionale
- Sector fitness: l'adozione per settore a livello regionale
- Entrambe calcolate come rapporto tra valore soluzione e indagine  $[0, 10]$
- Homogeneity fitness: coefficiente di variazione tra i tassi di adozione media comunale, se  $< 1$  calcolo  $(1 - c)*10$   $[0, 10]$

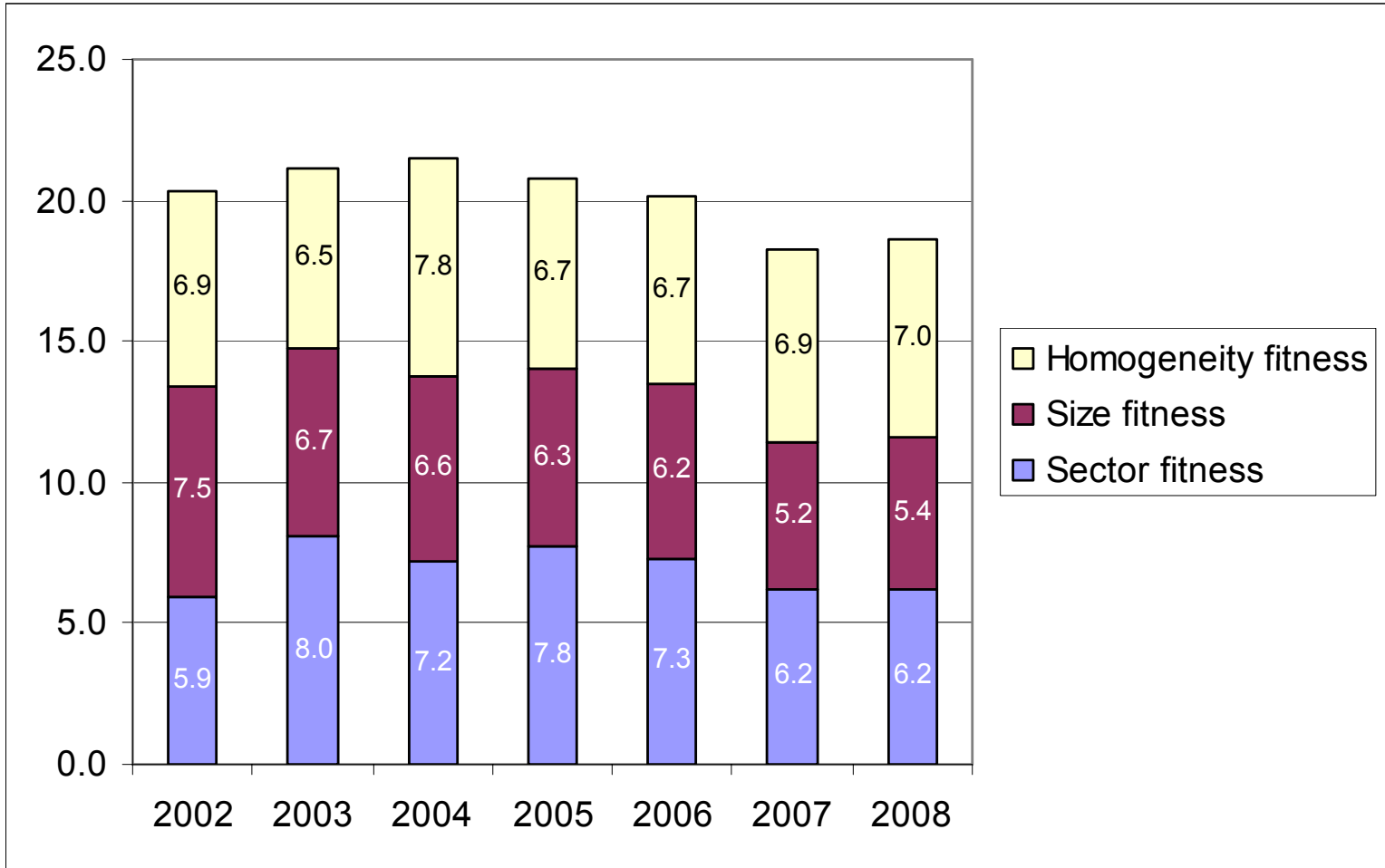


# Problemi di esecuzione

- Circa 30 ore di esecuzione su Server top causa operazioni di calcolo e necessità di RAM (circa 1,5 giga)
- Limitazioni al numero di soluzioni nella popolazione

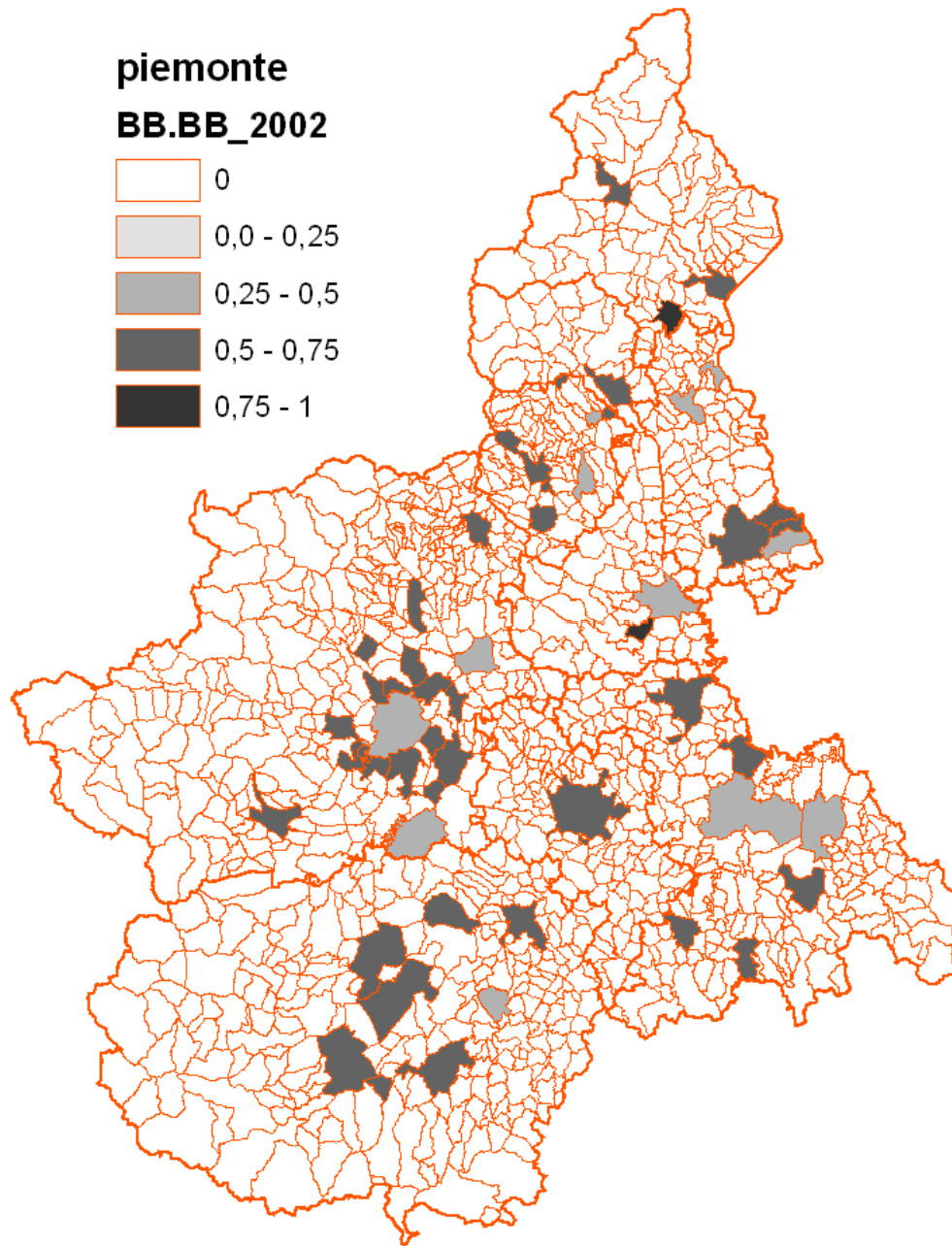
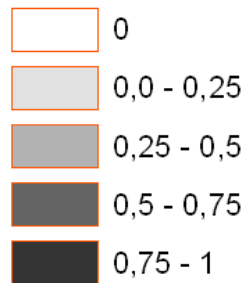
# Parametri utilizzati

- populationSize "200"
- stoppingMethod "2"
- stoppingValue "20000"
- normalizeFitness "1"
- mutateAdults "0"
- fittestNeverDie "1"
- turnoverRate "0.6"
- crossOverPointNumber "1"
- crossOverRate "0.9"
- mutationRate "0.05"



**piemonte**

**BB.BB\_2002**



# In conclusione

- L'aggiunta di informazioni (vincoli) a livello micro e sotto il micro è la chiave per la riuscita della stima
- Non è perfetto e nemmeno le indagini lo sono, ma si può fare