Walk Theory

*With Applications*

P.-L. Giscard, S. Thwaite, D. Jaksch

*Seminar*

15th August 2014
Outline

1. Introduction
   - Why walks?
   - Strange Observations

2. Walk Theory: the Math
   - Prime structure
   - Posets of walks and $\omega$-walks
   - Prime characterization of graphs

3. Case Study: Media-Industries
   - The network of media-industries
   - Verizon

4. Final Message
Why Walks?

Walk: a \textit{trajectory} on a graph

- Why walks?

Why Walks?

Walk: a *trajectory* on a graph

- Why walks?
  - Random walks, quantum random walks
  - Network analysis is often walk-based
  - Processes undergone by physical systems

Why Walks?

Walks are pervasive objects!

• Adjacency matrix $A^n =$ number of walks on graph
Why Walks?

Walks are pervasive objects!

- Adjacency matrix $A^n = \text{number of walks on graph}$
- Arbitrary matrix $M^n = \text{sum of walk weights}$

$$M = \begin{pmatrix} 1 & i \\ -3 & 5 \end{pmatrix}$$

Matrix power series are walk-series

Analytic matrix function $f(M) = \text{series of walk weights}$
Strange Observation 1

- Changing a square lattice

\[
\# W_{\bullet \rightarrow \bullet'}(\ell) \rightarrow \# W_{\bullet \rightarrow \bullet'}(2\ell)
\]
Strange Observation 1

- Changing a square lattice

\[
\# W_{\bullet \rightarrow \bullet'}(\ell) \quad \longrightarrow \quad \# W_{\bullet \rightarrow \bullet'}(2\ell)
\]

\[\leftarrow \text{Non-trivial for graphs: regularity is lost}\]
\[\leftarrow \text{Trivial transformation for walks}\]
Strange Observation 2

Which graphs are “similar”?

Which graphs are “similar”? 

Graph 1

Graph 2

Graph 3

Graph 4
Strange Observation 2

Which graphs are “similar”?

Yet the walk sets \( \mathcal{W}_{\cdot \rightarrow \cdot} \) are \textit{isomorphic} (will come back to that)
Strange Observation 3

- Why is network analysis difficult?

Example: characterizing molecules

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random walk kernel</td>
<td>Mutag(labelled)</td>
<td>90.0%</td>
</tr>
<tr>
<td>Backtrackless walk kernel</td>
<td>Mutag(labelled)</td>
<td>91.1%</td>
</tr>
<tr>
<td>Feature vector from Random walk</td>
<td>COIL(unlabeled)</td>
<td>94.4%</td>
</tr>
<tr>
<td>Feature vector from backtrackless random walk</td>
<td>COIL(unlabeled)</td>
<td>95.5%</td>
</tr>
<tr>
<td>Feature vector from Ihara coefficients</td>
<td>COIL(unlabeled)</td>
<td>94.4%</td>
</tr>
<tr>
<td>Shortest Path Kernel</td>
<td>COIL(unlabeled)</td>
<td>86.7%</td>
</tr>
<tr>
<td>Feature vector from Random walk</td>
<td>Mutag(unlabeled)</td>
<td>89.4%</td>
</tr>
<tr>
<td>Feature vector from backtrackless random walk</td>
<td>Mutag(unlabeled)</td>
<td>90.5%</td>
</tr>
<tr>
<td>Feature vector from Ihara coefficients</td>
<td>Mutag(unlabeled)</td>
<td>80.5%</td>
</tr>
</tbody>
</table>


Walk-based methods
Strange Observations

- Completely *dissimilar* graphs can have the *same* walk sets
- Even *fundamental* graph properties, *regularity*, may have little effect on walks
- Existence of *non-trivial* properties of walks *valid on all* multi-digraphs
- More strange objects to come (*walks without graphs!*)

Why should we care?
Danger

- The *danger* of misunderstanding walks

Which vertex is most central?
Danger

- The *danger* of misunderstanding walks

Which vertex is most central?

Both red vertices have the same centrality...
...worse, all *self-communicability* measures are identical
Elements of Walk Theory

1. Introduction
   - Why walks?
   - Strange Observations

2. Walk Theory: the Math
   - Prime structure
   - Posets of walks and $\omega$-walks
   - Prime characterization of graphs

3. Case Study: Media-Industries
   - The network of media-industries
   - Verizon

4. Final Message
Prime Structure

Observation: walk = simple path & simple cycle

Walk $w$ factors into 1 simple path and 2 simple cycles

$$w = \gamma_0 \circ \gamma_1 \circ \gamma_2$$

What can we say in general?
Theorem

Let \( \mathcal{G} \) be a graph and \( w \) a walk on \( \mathcal{G} \).
Then there exists a unique factorisation of \( w \) into \( \odot \)-products of prime walks, the simple paths and simple cycles on \( \mathcal{G} \).

Prime walks: \( \gamma \mid w \odot w' \Rightarrow \gamma \mid w \) or \( \gamma \mid w' \)

This holds on all multi-digraphs!

Walk factorization is efficient \( t \propto O(\ell_w) \)

\[
1232332121 = \left( 121 \odot \left( 232 \odot \left( 232 \odot 33 \right) \right) \right) \odot 121
\]

Primes provide walk sets with a structure
Posets of Walks

Construct a prime-based representation of walks

- Sets of walks ordered by divisibility

\[ w \preceq \odot w' \iff w \mid w' \]

Example: \( 121 \mid 1221 = 121 \odot 22 \implies 121 \preceq \odot 1221 \)

- Poset of walks \( P_\alpha = (W_G; \alpha \alpha, \preceq \odot) \)

Set of walks from \( \alpha \) to \( \alpha \) on \( G \)
Posets of Walks

Construct a prime-based representation of walks

- Sets of walks ordered by divisibility

\[ w \leq_{\circ} w' \iff w | w' \]

Example: \( 121 | 1221 = 121 \circ 22 \implies 121 \leq_{\circ} 1221 \)

- Poset of walks \( P_{\alpha} = (W_G; \alpha \alpha, \leq_{\circ}) \)

Set of walks from \( \alpha \) to \( \alpha \) on \( G \)

\( P_{\alpha} \) is a complicated object: \( \infty \)-many walks

Difficult to find all divisors

Can we make things simpler?
Posets of $\omega$-walks

Sets of distinct prime factors...

$$w = \gamma_0 \odot \gamma_1^3 \Rightarrow S_\omega(w) = \{\gamma_0, \gamma_1\}$$

...ordered by inclusion

$$S_\omega(w) \leq_\omega S_\omega(w') \iff S_\omega(w) \subseteq S_\omega(w')$$

Poset of $\omega$-walks: $P_\omega = (S_\omega, \leq_\omega)$

An element of $P_\omega$ is a list of distinct cycles
Example: $P_\alpha^\omega$ on the Square

Walking from • to itself
Which primes can be reached?

$\rightarrow$ Set of accessible primes, walking from • to itself
Example: $P_\omega^\alpha$ on the Square

Relations between the primes
Example: $P^\omega_\alpha$ on the Square

Relations between the primes

Prime Tree $T^\omega$
Example: $P^\omega_\alpha$ on the Square

Primes provide walk sets with a *structure*

= $P^\omega$
Posets of $\omega$-walks

$P^\omega_\alpha$ is simpler than $P_\alpha$

- $P^\omega_\alpha$ is a finite lattice, has a smallest and largest element
- $P^\omega_\alpha$ is the uniquely determined by the prime tree $T^\omega_\alpha$

Theorem
Knowledge of $P^\omega_\alpha$ or $T^\omega_\alpha$ is equivalent to that of $P_\alpha$.

- Number of walk posets on $n$ primes:
  \[ B_n \sim \left(\frac{.792}{\log n} n\right)^n \]
Strange Observation 4

- Number of walk posets on $n$ primes:
  
  $$B_n \sim \left(\frac{.792 \, n}{\log n}\right)^n$$

- Most walk sets have no exact graph realization!
Strange Observation 4

- Number of walk posets on $n$ primes:
  \[ B_n \sim \left( \frac{.792 \, n}{\log n} \right)^n \]
Strange Observation 4

- Number of walk posets on \( n \) primes:
  \[ B_n \sim \left(0.792 \frac{n}{\log n}\right)^n \]

- Number of walk posets on \( n \) primes with exact graph realization
  \[ C_n \sim 4^n n^{-3/2} / \sqrt{\pi} \ll B_n \]

\( \subseteq P^\omega \)

No graph exists with exactly this walk poset

\( a, b, c, d \) primes

\( \leftarrow \) Most walk sets have no exact graph realization!
Isomorphic Walk Sets

Isomorphic walk sets on dissimilar graphs

**Theorem**

Let $G_1$ and $G_2$ be two multi-digraphs and $\alpha$ and $\alpha$ two vertices on $G_1$ and $G_2$, respectively. If $T_{G_1;\alpha}$ is isomorphic to $T_{G_2;a}$ then $(W_{G_1;\alpha \alpha}, \circ)$ is isomorphic to $(W_{G_2;aa}, \circ)$.

Length of prime $\ell_p$ is *unspecified*

$\rightarrow \infty$-many graph realizations
Prime Characterization of Graphs

Graphs determine walks...  *do walks determine graphs?*

**YES**

**Theorem**

Let $G$ be a connected multi-digraph.  
Then $G$ is *uniquely* determined, up to an *isomorphism*, by the primes it sustains.

$\leftrightarrow$ Prime trees + length of primes + roots $=$ unique graph

↑ location of the 1$^{\text{st}}$ vertex
Elements of Walk Theory

1. Introduction
   - Why walks?
   - Strange Observations

2. Walk Theory: the Math
   - Prime structure
   - Posets of walks and $\omega$-walks
   - Prime characterization of graphs

3. Case Study: Media-Industries
   - The network of media-industries
   - Verizon

4. Final Message
Network of Media-Industries

Industries linked by thick edges & feedback processes

Source: Media Industry Networks 2006, Laurie Lock Lee, Optimice
Network of Media-Industries

Prime trees

Universal and perfect representation of walk sets $W_G$
Network of Media-Industries

→ "Distance" between prime trees:
Verizon vs Yahoo

- Network analysis
  Centrality; Eigenvector centr.: Yahoo > Verizon
  Degree centr.; Closeness centr.: Yahoo > Verizon

- Distinct feedback processes
  Verizon 100 > Yahoo 55

- Trees of cycles: $1 \leq (\text{DFP} / \# \text{primes}) \leq \text{Comp. gr. of cycles}$
  Verizon 4.5 > Yahoo 2.9
Verizon vs Microsoft

- Network analysis
  Centrality; Eigenvector centr. : VeriZon > Microsoft
  Degree centr.; Closeness centr.: VeriZon > Microsoft

- Distinct feedback processes: VeriZon 100 ≳ Microsoft 95
- (DFP/ # primes): VeriZon 4.5 ≳ Microsoft 4.3
- Participation to feedbacks: VeriZon = Microsoft
Verizon vs Time Warner

- Network analysis
  Centrality; Eigenvector centr. : **VeriZon** > Time Warner

- Distinct feedback processes: **VeriZon** 100 > Time Warner 87
- (DFP/ # primes): **VeriZon** 4.5 > Time Warner 3.95
- Participation to feedbacks: **VeriZon** < Time Warner

\[ \text{Time Warner participates more than Verizon?} \]
Return to equilibrium

Let's perturb Verizon and Time Warner...

Time Warner has more impact than Verizon!
Walk Theory

1. Introduction
   - Why walks?
   - Strange Observations

2. Walk Theory: the Math
   - Prime structure
   - Posets of walks and $\omega$-walks
   - Prime characterization of graphs

3. Case Study: Media-Industries
   - The network of media-industries
   - Verizon

4. Final Message
Final Message

Are you interested in the processes that a system undergoes or the network that sustains them?

Main Message

*Walks need to be treated by a separate theory of walks. It provides new graph-free tools to study them.*

Results

- Graph-free representation of walks using primes
- Rigorous mathematical theory
- Enables perfect comparison of walk sets

Open problems

- There is much more to say about prime trees!
Thank You!

**Work supported by: EPSRC Grant EP/K038311/1**


I am looking for a postdoctoral position!