Semirings for Social Networks Analysis

Vladimir Batagelj *
Department of Mathematics, University of Ljubljana
Jadranska 19, 61 111 Ljubljana, Slovenia
E–mail: VLADIMIR.BATAGELJ@UNI-LJ.SI

Abstract

In the paper four semirings for solving social networks problems are constructed.

The closures of the matrix of a given signed graph over balance and cluster semirings can be used to decide whether the graph is balanced or clusterable.

The closure of relational matrix over geodetic semiring contains for every pair of vertices u and v the length and the number of u-v geodesics; and for geosetic semiring the length and the set of vertices on u-v geodesics. The algorithms for computing the geodetic and the geosetic closure matrix are also given.

Key words : balanced signed graphs, clusterable signed graphs, closed semirings, closure, geodesic, Freeman's centrality indices, Boyle's operation.

Math. Subj. Class. (1991) : 05 C 50, 05 C 12, 05 C 75, 16 Y 60, 68 R 10, 92 H 30

There are several applications of semirings in social networks analysis. For example, the semiring $(\{0, 1, 2, 3\}, \max, \min)$ can be applied to determine the connectedness matrix [19, p. 133]. Similar semiring $(\{0, u, m\}, \max, \min)$, $0 < u < m$ (0 – no link, u – uniplex links, m – multiplex links) can be used to analyse the connectivity of social networks [12].

In this paper we shall construct four new semirings for solving social networks analysis problems:

- *balance* and *cluster* semirings which can be used to decide whether the graph is balanced or clusterable;
- *geodetic* semiring to determine for every pair of vertices u and v the length and the number of u-v geodesics;
- *geosetic* semiring for computing Boyle's operation – to determine for every pair of vertices u and v the length and the set of vertices on u-v geodesics.

*Supported in part by the Ministry of Science and Technology of Slovenia.

1 Semirings

An algebraic structure $(A, +, \cdot, 0, 1)$ on the set A is a *semiring* [1, 5, 6, 22] iff:

- $(A, +, 0)$ is an abelian monoid with neutral element 0 (*zero*);
- $(A, \cdot, 1)$ is a monoid with neutral element 1 (*unit*);
- multiplication $\cdot$ distributes over addition $+$; for all $a, b, c \in A$

  $a \cdot (b + c) = a \cdot b + a \cdot c \quad$ and $\quad (a + b) \cdot c = a \cdot c + b \cdot c$

  In the expressions we assume precedence of multiplication over addition.

The semiring $(A, +, \cdot, 0, 1)$ is *complete* iff the addition is well-defined also for countable sets of elements and the (generalized) commutativity, associativity (for addition) and distributivity hold also in this case.

If the addition is *idempotent*, for every $a \in A : a + a = a$, the semiring over a finite set A is complete.

The semiring $(A, +, \cdot, ^\star, 0, 1)$ is *closed* iff for the (unary) *closure* operation * it holds for every $a \in A$

$$a^\star = 1 + a \cdot a^\star = 1 + a^\star \cdot a$$

There can exist different closures over the same semiring.

The complete semiring is always closed for the closure defined by

$$a^\star = \sum_{k=0}^{\infty} a^k$$

In the sequel we shall refer by the term *closure* to the operation defined by this expression. In a closed semiring we can define also a *strict closure* $\overline{a}$ by

$$\overline{a} = a \cdot a^\star$$

Suppose that for a given graph $G = (V, R),\ R \subseteq V \times V$ and a semiring $(A, +, \cdot, 0, 1)$ a *value function*

$$d : R \to A$$

is given.

A finite sequence of vertices $\pi = v_0, v_1, v_2, \ldots, v_{p-1}, v_p$ is a *walk* of *length* p on G iff $v_{i-1} R v_i, i = 1, \ldots, p$; and is a *semiwalk* or *chain* on G iff $v_{i-1} R v_i \vee v_i R v_{i-1}, i = 1, \ldots, p$. The (semi)walk is *closed* iff its ends coincide, $v_0 = v_p$. A walk in which no vertex appears twice is called a *path*; and if only its ends coincide it is called a *cycle*.

We can extend the value function d to walks and sets of walks on G by

- let ε_v be a null walk in the vertex $v \in V$ then $d(\varepsilon_v) = 1$
- let $\pi = v_0, v_1, v_2, \ldots, v_{p-1}, v_p$ be a walk of length $p \geq 1$ on G then

  $d(\pi) = d(v_0, v_1) \cdot d(v_1, v_2) \ldots d(v_{p-1}, v_p)$
- for the empty set of walks $\emptyset$ we have $d(\emptyset) = 0$

- let $\mathcal{P} = \{\pi_1, \pi_2, \ldots\}$ be a set of walks on G then
  $d(\mathcal{P}) = d(\pi_1) + d(\pi_2) + \ldots$

We denote by $\mathcal{P}^p_{uv}$ the set of all walks of length p from vertex u to vertex v; by $\mathcal{P}^\star_{uv}$ the set of all walks from vertex u to vertex v; and by $\overline{\mathcal{P}}_{uv}$ the set of all nontrivial (different from ε_u) walks from vertex u to vertex v.

The *value matrix* of a graph is a matrix $\mathbf{D}$ defined by

$$\mathbf{D}[u, v] = \begin{cases} d((u, v)) & (u, v) \in R \\ 0 & \text{otherwise} \end{cases}$$

In the following we shall assume that in the semiring for every $a \in A$

$$a \cdot 0 = 0 \cdot a = 0$$

holds and that the set of vertices is finite $V = \{v_1, v_2, \ldots, v_n\}$. Then the addition and multiplication can be extended in the usual way to square matrices of order n which themselves form a semiring.

The matrix semiring over complete semiring is also complete and therefore closed for

$$\mathbf{D}^\star = \sum_{k=0}^{\infty} \mathbf{D}^k$$

There are two well known theorems [1, 6, 22] connecting values of walks in graphs and their matrices:

THEOREM 1. Let $\mathbf{D}^p$ be the p-th power of value matrix $\mathbf{D}$ then

$$d(\mathcal{P}^p_{uv}) = \mathbf{D}^p[u, v]$$

THEOREM 2. Let $\mathbf{D}$ be a value matrix over complete semiring, $\mathbf{D}^\star$ its closure and $\overline{\mathbf{D}}$ its strict closure matrix then

$$d(\mathcal{P}^\star_{uv}) = \mathbf{D}^\star[u, v] \qquad \text{and} \qquad d(\overline{\mathcal{P}}_{uv}) = \overline{\mathbf{D}}[u, v]$$

To compute the closure matrix $\mathbf{D}^\star$ of a given matrix $\mathbf{D}$ over a complete semiring $(A, +, \cdot, 0, 1, ^\star)$ we can use the Fletcher's algorithm [14]:

$\mathbf{C}_0 := \mathbf{D}$;
for $k := 1$ **to** n **do begin**
  for $i := 1$ **to** n **do for** $j := 1$ **to** n **do**
    $c_k[i, j] := c_{k-1}[i, j] + c_{k-1}[i, k] \cdot (c_{k-1}[k, k])^\star \cdot c_{k-1}[k, j]$;
  $c_k[k, k] := 1 + c_k[k, k]$;
end;
$\mathbf{D}^\star := \mathbf{C}_n$;

If we delete the statement $c_k[k, k] := 1 + c_k[k, k]$ we obtain the algorithm for computing the strict closure $\overline{\mathbf{D}}$.

If the addition is idempotent we can by this algorithm compute the closure matrix in place – we omit the subscripts in matrices $\mathbf{C}$.

There is a bijection between semiwalks on graph G and walks in the graph $\hat{G}$ which is the sum of the graph G and its inverse – graph G with reversed directions of its arcs. We have

$$\mathbf{D}(\hat{G}) = \mathbf{D}(G) + \mathbf{D}(G)^T$$

where $\mathbf{D}^T$ denotes the transpose of matrix $\mathbf{D}$. Let us define the *symmetric* closure of the value matrix by

$$\mathbf{D}^\bullet = (\mathbf{D} + \mathbf{D}^T)^\star$$

2 Balance and cluster semirings

2.1 Balanced and clusterable signed graphs

A *signed graph* is an ordered pair (G, σ) where

- $G = (V, R)$ is a directed graph (without loops) with set of vertices V and set of arcs $R \subseteq V \times V$;
- $\sigma : R \to \{p, n\}$ is a *sign* function. The arcs with the sign p are *positive* and the arcs with the sign n are *negative*. We denote the set of all positive arcs by R^+ and the set of all negative arcs by R^-.

The case when the graph is undirected can be reduced to the case of directed graph by replacing each edge e by a pair of opposite arcs both signed with the sign of the edge e.

The signed graphs were introduced in [18] and later studied by several authors [7, 8, 10, 11, 19, 20, 21]. Following Roberts [21, pp. 75–77] a signed graph (G, σ) is:

- *balanced* iff the set of vertices V can be partitioned into two subsets so that every positive arc joins vertices of the same subset and every negative arc joins vertices of different subsets.
- *clusterable* iff the set of V can be partitioned into subsets, called *clusters*, so that every positive arc joins vertices of the same subset and every negative arc joins vertices of different subsets.

The (semi)walk on the signed graph is *positive* iff it contains an even number of negative arcs; otherwise it is *negative*.

The balanced and clusterable signed graphs are characterised by the following theorems [9, 10, 19, 20, 21].

THEOREM 3. A signed graph (G, σ) is balanced iff every closed semiwalk is positive.

THEOREM 4. A signed graph (G, σ) is clusterable iff G contains no closed semiwalk with exactly one negative arc.

Table 1: Balance semiring

$+$	0	n	p	a
0	0	n	p	a
n	n	n	a	a
p	p	a	p	a
a	a	a	a	a

$\cdot$	0	n	p	a
0	0	0	0	0
n	0	p	n	a
p	0	n	p	a
a	0	a	a	a

x	$x^\star$
0	p
n	a
p	p
a	a

Table 2: Cluster semiring

$+$	0	n	p	a	q
0	0	n	p	a	q
n	n	n	a	a	n
p	p	a	p	a	p
a	a	a	a	a	a
q	q	n	p	a	q

$\cdot$	0	n	p	a	q
0	0	0	0	0	0
n	0	q	n	n	q
p	0	n	p	a	q
a	0	n	a	a	q
q	0	q	q	q	q

x	$x^\star$
0	p
n	a
p	p
a	a
q	p

2.2 Balance and cluster semirings

To construct a semiring corresponding to the *balance* problem we take the set A with four elements [11, 19]:

- 0 no walk;
- n all walks are negative;
- p all walks are positive;
- a at least one positive and at least one negative walk.

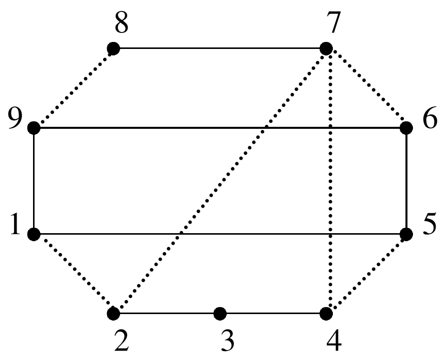
Now it is easy to produce the Cayley tables for *balance semiring* (see Table 1). The balance semiring is idempotent closed semiring with zero 0 and unit p.

For construction of the *cluster semiring* corresponding to the *clusterability* problem we need the set A with five elements:

- 0 no walk;
- n at least one walk with exactly one negative arc;
no walk with only positive arcs;
- p at least one walk with only positive arcs;
no walk with exactly one negative arc;
- a at least one walk with only positive arcs;
at least one walk with exactly one negative arc;
- q each walk has at least two negative arcs.

The Cayley tables for *cluster semiring* are given in Table 2. The cluster semiring is idempotent closed semiring with zero 0 and unit p.

Combining the Theorem 2 with Theorems 3 and 4 we get.



	1	2	3	4	5	6	7	8	9
1	0	n	0	0	p	0	0	0	p
2	n	0	p	0	0	0	n	0	0
3	0	p	0	p	0	0	0	0	0
4	0	0	p	0	n	0	n	0	0
5	p	0	0	n	0	p	0	0	0
6	0	0	0	0	p	0	n	0	p
7	0	n	0	n	0	n	0	p	0
8	0	0	0	0	0	0	p	0	n
9	p	0	0	0	0	p	0	n	0

Figure 1: Chartrand's example – graph

THEOREM 3'. A signed graph (G, σ) is balanced iff the diagonal of its balance-closure matrix $\mathbf{D}_B^\bullet$ contains only elements with value p.

THEOREM 4'. A signed graph (G, σ) is clusterable iff the diagonal of its cluster-closure matrix $\mathbf{D}_C^\bullet$ contains only elements with value p.

The balance-closure matrix of balanced signed graph contains no element with value a, since in this case the corresponding diagonal elements should also have value a. Similary the cluster-closure matrix of clusterable signed graph contains no element with value a.

A block is a maximal set of vertices with equal lines in matrix $\mathbf{D}^\bullet$.

In balance-closure of balanced signed graph and in cluster-closure of clusterable signed graph all the entries between vertices of two blocks have the same value. The value of entries between vertices of the same block is p.

In both cases different partitions of the set of vertices correspond to the (nonequivalent) colorings of the graph with blocks as vertices in which there is an edge between two vertices iff the entries between the corresponding blocks in matrix $\mathbf{D}^\bullet$ have value n.

There is another way to test the clusterability of a given signed graph:

THEOREM 4''. A signed graph (G, σ) is clusterable iff $(R^+)^\bullet \cap R^- = \emptyset$, where the closure $^\bullet$ is computed in the semiring $(\{0, 1\}, \vee, \wedge, 0, 1)$.

This form of the Theorem 4 is interesting because the intersection $(R^+)^\bullet \cap R^-$ consists of arcs which prevent the signed graph (G, σ) to be clusterable.

2.3 Examples

EXAMPLE 1. In Figure 1 the graph from [9, page 181] and its value matrix are given. The positive edges are drawn with solid lines, and the negative edges with dotted lines.

	1	2	3	4	5	6	7	8	9
1	a	a	a	a	a	a	a	a	a
2	a	a	a	a	a	a	a	a	a
3	a	a	a	a	a	a	a	a	a
4	a	a	a	a	a	a	a	a	a
5	a	a	a	a	a	a	a	a	a
6	a	a	a	a	a	a	a	a	a
7	a	a	a	a	a	a	a	a	a
8	a	a	a	a	a	a	a	a	a
9	a	a	a	a	a	a	a	a	a

	1	2	3	4	5	6	7	8	9
1	**p**	n	n	n	**p**	**p**	n	n	**p**
2	n	**p**	**p**	**p**	n	n	n	n	n
3	n	**p**	**p**	**p**	n	n	n	n	n
4	n	**p**	**p**	**p**	n	n	n	n	n
5	**p**	n	n	n	**p**	**p**	n	n	**p**
6	**p**	n	n	n	**p**	**p**	n	n	**p**
7	n	n	n	n	n	n	**p**	**p**	n
8	n	n	n	n	n	n	**p**	**p**	n
9	**p**	n	n	n	**p**	**p**	n	n	**p**

Table 3: Chartrand's example – closures

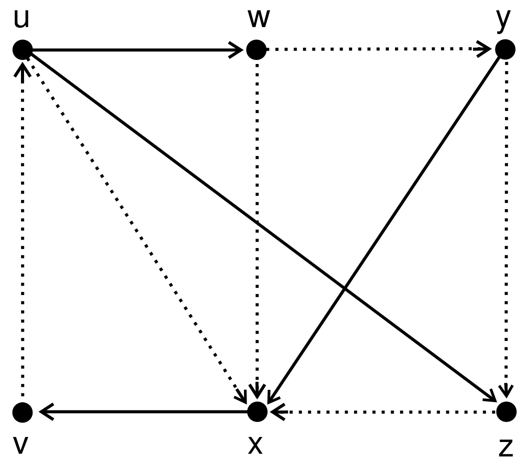


Figure 2: Roberts's example – graph

On the left side of Table 3 the corresponding balance-closure is given – the graph is not balanced. From the cluster-closure on the right side of Table 3 we can see that the graph is clusterable and it has the clusters

$$V_1 = \{1, 5, 6, 9\}, \quad V_2 = \{2, 3, 4\}, \quad V_3 = \{7, 8\}$$

EXAMPLE 2. The signed graph from [21, page 77, exercise 16] is presented in Figure 2, and its value matrix on the left side of Table 4. In this case the balance-closure and the cluster-closure are equal (right side of Table 4). The corresponding partition is

$$V_1 = \{v, x, y\}, \quad V_2 = \{u, w, z\}$$

	u	v	w	x	y	z
u	0	0	p	n	0	p
v	n	0	0	0	0	0
w	0	0	0	n	n	0
x	0	p	0	0	0	0
y	0	0	0	p	0	n
z	0	0	0	n	0	0

	u	v	w	x	y	z
u	**p**	n	**p**	n	n	**p**
v	n	**p**	n	**p**	**p**	n
w	**p**	n	**p**	n	n	**p**
x	n	**p**	n	**p**	**p**	n
y	n	**p**	n	**p**	**p**	n
z	**p**	n	**p**	n	n	**p**

Table 4: Roberts's example – value matrix and its closure

3 Geodetic semiring

3.1 Construction of geodetic semiring

Another example of a semiring we found reading the book [17, p. 34-38, 111-112]. In 1977 Freeman introduced the centrality index of a vertex based on betweenness [15, 16]:

$$C_B(t) = \sum_u \sum_v \frac{n_{u,v}(t)}{n_{u,v}}$$

where $n_{u,v}$ is the number of geodesics from vertex u to vertex v; and $n_{u,v}(t)$ is the number of geodesics from u to v that contain vertex t.

For computing $C_B(t)$ he proposed the methods given in Harary, Norman, Cartwright [19, p. 134-141].

Here we present an alternative approach to computing $C_B(t)$. Suppose that we know a matrix

$$[(d_{u,v}, n_{u,v})]$$

where $d_{u,v}$ is the length of u-v geodesics and $n_{u,v}$ is the number of u-v geodesics. Then it is also easy to determine $n_{u,v}(t)$:

$$n_{u,v}(t) = \begin{cases} n_{u,t} \cdot n_{t,v} & d_{u,t} + d_{t,v} = d_{u,v} \\ 0 & \text{otherwise} \end{cases}$$

Matrix $[(d_{u,v}, n_{u,v})]$ can be obtained by computing the closure of relation matrix over the following *geodetic semiring*.

First we transform relation R to a matrix $\mathbf{R} = [(d,n)_{u,v}]$ which has for entries pairs defined by

$$(d,n)_{u,v} = \begin{cases} (1,1) & (u,v) \in R \\ (\infty, 0) & (u,v) \notin R \end{cases}$$

where d is the length of a shortest path and n is the number of shortest paths.

In the set $A = (\mathbb{R}_0^+ \cup \{\infty\}) \times (\mathbb{N} \cup \{\infty\})$ we define two operations: *addition*:

$$(a,i) \oplus (b,j) = (\min(a,b), \begin{cases} i & a < b \\ i+j & a = b \\ j & a > b \end{cases})$$

and *multiplication*:

$$(a,i) \odot (b,j) = (a+b, i \cdot j)$$

It is easy to verify that $(A, \oplus, \odot)$ is indeed a semiring with zero $(\infty, 0)$ and identity $(0,1)$. The verifications of semiring properties are straightforward. We present only the less trivial verifications of associativity of addition and of distributivity:

Associativity:

$$((a,i) \oplus (b,j)) \oplus (c,k) = (a,i) \oplus ((b,j) \oplus (c,k))$$

Since the addition $\oplus$ is commutative we can assume that $a \leq b$.

case	$((a,i) \oplus (b,j)) \oplus (c,k)$	$(a,i) \oplus ((b,j) \oplus (c,k))$
$a < b, b < c \Rightarrow a < c$	$(a,i) \oplus (c,k) = (a,i)$	$(a,i) \oplus (b,j) = (a,i)$
$a < b, b = c \Rightarrow a < c$	$(a,i) \oplus (c,k) = (a,i)$	$(a,i) \oplus (b,j+k) = (a,i)$
$a < b, b > c$	$(a,i) \oplus (c,k)$	$(a,i) \oplus (c,k)$
$a = b, b < c \Rightarrow a < c$	$(a,i+j) \oplus (c,k) = (a,i+j)$	$(a,i) \oplus (b,j) = (a,i+j)$
$a = b, b = c$	$(a,i+j) \oplus (c,k) = (a,i+j+k)$	$(a,i) \oplus (b,j+k) = (a,i+j+k)$
$a = b, b > c$	$(a,i+j) \oplus (c,k) = (c,k)$	$(a,i) \oplus (c,k) = (c,k)$

From the table we can see that in all possible cases the value of the left side term equals the value of the right side term. Therefore the addition $\oplus$ is associative.

Distributivity:

$$(a,i) \odot ((b,j) \oplus (c,k)) = (a,i) \odot (b,j) \oplus (a,i) \odot (c,k)$$

Because of the commutativity of the addition $\oplus$, we can assume that $b \leq c$. We obtain for the left side term

$$(a,i) \odot ((b,j) \oplus (c,k)) = (a,i) \odot (b, \begin{cases} j & b < c \\ j+k & b = c \end{cases}) = (a+b, i \cdot \begin{cases} j & b < c \\ j+k & b = c \end{cases})$$

and for the right side term

$$(a,i) \odot (b,j) \oplus (a,i) \odot (c,k) = (a+b, i \cdot j) \oplus (a+c, i \cdot k) = (a+b, \begin{cases} i \cdot j & b < c \\ i \cdot j + i \cdot k & b = c \end{cases})$$

We get the same value on both sides. Therefore the distributivity holds.

The semiring $(A, \oplus, \odot)$ is also complete and closed, with a closure

$$(a,i)^\star = \begin{cases} (0,\infty) & a = 0, i \neq 0 \\ (0,1) & \text{otherwise} \end{cases}$$

This can be easily verified

$$\begin{aligned} (0,1) \oplus (a,i) \odot (a,i)^\star &= (0,1) \oplus (a + a_\star, i \cdot i_\star) \\ &= (0,1) \oplus \begin{cases} (0,\infty) & a + a_\star = 0, i \neq 0 \\ (0,0) & a + a_\star = 0, i = 0 \\ (a + a_\star, i \cdot i_\star) & a + a_\star > 0 \end{cases} \\ &= \begin{cases} (0,\infty) & a = 0, i \neq 0 \\ (0,1) & \text{otherwise} \end{cases} = (a,i)^\star \end{aligned}$$

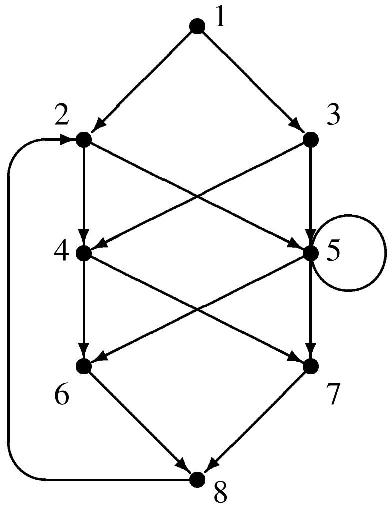


Figure 3: Example graph

	1	2	3	4	5	6	7	8
1	$(\infty, 0)$	$(1, 1)$	$(1, 1)$	$(2, 2)$	$(2, 2)$	$(3, 4)$	$(3, 4)$	$(4, 8)$
2	$(\infty, 0)$	$(4, 4)$	$(\infty, 0)$	$(1, 1)$	$(1, 1)$	$(2, 2)$	$(2, 2)$	$(3, 4)$
3	$(\infty, 0)$	$(4, 4)$	$(\infty, 0)$	$(1, 1)$	$(1, 1)$	$(2, 2)$	$(2, 2)$	$(3, 4)$
4	$(\infty, 0)$	$(3, 2)$	$(\infty, 0)$	$(4, 2)$	$(4, 2)$	$(1, 1)$	$(1, 1)$	$(2, 2)$
5	$(\infty, 0)$	$(3, 2)$	$(\infty, 0)$	$(4, 2)$	$(1, 1)$	$(1, 1)$	$(1, 1)$	$(2, 2)$
6	$(\infty, 0)$	$(2, 1)$	$(\infty, 0)$	$(3, 1)$	$(3, 1)$	$(4, 2)$	$(4, 2)$	$(1, 1)$
7	$(\infty, 0)$	$(2, 1)$	$(\infty, 0)$	$(3, 1)$	$(3, 1)$	$(4, 2)$	$(4, 2)$	$(1, 1)$
8	$(\infty, 0)$	$(1, 1)$	$(\infty, 0)$	$(2, 1)$	$(2, 1)$	$(3, 2)$	$(3, 2)$	$(4, 4)$

Table 5: Geodetic closure

A semiring element (a, i) is *positive* iff $a > 0$. The set A^+ of all positive elements is closed for addition and multiplication. Note also that for positive elements the absorbtion property

$$(0, 1) \oplus (a, i) = (0, 1)$$

holds.

3.2 Algorithm and example

Let $(d, c)_{u,v}$ be the entry of the strict closure $\overline{\mathbf{R}}$ of relation matrix **R** over geodetic semiring. Then d equals to the length of a shortest nontrivial u-v path (a geodesic or a shortest cycle), and c equals to the number of different u-v geodesics.

Using this algorithm we obtained for a graph represented in Figure 3. the strict geodetic closure presented in Table 5.

To adapt the Fletcher's algorithm for computing the strict geodetic closure we have to consider some properties of geodetic semiring. By the construction all entries of the relation matrix **R** are positive. From the description of Fletcher's algorithm it follows that also the entry $c_{k-1}[k,k]$ in $(c_{k-1}[k,k])^\star$ is always positive. Therefore $(c_{k-1}[k,k])^\star = (0,1)$ and we can omit it from the expression. A detailed analysis of the algorithm shows that we can compute the strict geodetic closure matrix in place – we omit the subscripts in matrices **C**.

Representing the relation matrix **R** (and its geodetic closure) by two matrices: the shortest paths length matrix **D** and the geodesics count matrix **C**, we obtain the following adapted version of Fletcher's algorithm for computing the strict geodetic closure:

```
for k := 1 to n do begin
    for i := 1 to n do for j := 1 to n do begin
        dst := min(big, d[i,k] + d[k,j]);
        if d[i,j] ≥ dst then begin
            cnt := c[i,k] * c[k,j];
            if d[i,j] = dst then c[i,j] := c[i,j] + cnt
                else begin c[i,j] := cnt; d[i,j] := dst end;
        end;
    end;
end;
```

The constant big in the algorithm is a number representing the infinity ∞.

4 Semiring for Boyle's operation

4.1 Construction of geosetic semiring

In [4, 13] Boyle introduced the operation

$$u * v = \text{set of vertices of a graph } G \text{ that belong to the shortest paths from } u \text{ to } v$$

over the set V of vertices of a graph $G = (V, R), \quad R \subseteq V \times V$.

This operation has several interesting properties. For example:

$$u * v \subset u * t \ \Rightarrow\ v * t \subset u * t,$$
$$u * v = w * t \ \Rightarrow\ v, u, w, t \text{ belong to the same cycle.}$$

But, it is not associative. Nevertheless it is possible to embed it into a semiring in the following way.

We start with *quadruples* of the form

$$(u, P, v, p)$$

with the interpretation: $u \in V$ – initial vertex, $v \in V$ – terminal vertex, $P \subseteq V$ – vertices on the shortest u–v paths, $p \in \mathbb{N}$ – length of these paths. We require also that $\{u, v\} \subseteq P$ and $p <$ card P. Quadruples with length 0 are of the form $(u, \{u\}, u, 0), u \in V$.

We build the *geosetic* semiring over the set $\mathcal{S}$ of finite sets of quadruples $\alpha \in \mathcal{S}$

$$\alpha = \{a_i\}, \quad a_i = (u_i, P_i, v_i, p_i)$$

which also satisfy the requirement that no pair of quadruples in α has the same initial and terminal vertices

$$u_i = u_j \wedge v_i = v_j \Rightarrow i = j$$

Let us first consider one element sets

$$\alpha = \{a\},\ a = (u, P, v, p) \quad \text{and} \quad \beta = \{b\},\ b = (w, Q, t, q)$$

We can define *addition* as follows:

$$\alpha \oplus \beta = \begin{cases} \left.\begin{array}{ll} \{(u, P \cup Q, v, p)\} & p = q \\ \alpha & p < q \\ \beta & q < p \end{array}\right\}, \ u = w,\ v = t \\ \alpha \cup \beta \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

Note that in the last case $\alpha \cup \beta = \{a, b\}$ which is not a one element set so that the addition is not closed over one element sets. To extend it to $\mathcal{S}$ we first introduce the operation of *reduction* $\text{Red}\,\alpha$, which from an arbitrary finite set of quadruples, α, eliminates multiple occurrences of quadruples with the same initial and terminal vertices. It is defined by

$$\text{Red}\,\alpha = [(\alpha \setminus \{a_i, a_j\}) \cup (\{a_i\} \oplus \{a_j\}), i < j]$$

or in a procedural form

> $\beta := \alpha;$
> **while** $\exists i, j : (i \neq j \wedge u_i = u_j \wedge v_i = v_j)$ **do** $\beta := (\beta \setminus \{a_i, a_j\}) \cup (\{a_i\} \oplus \{a_j\});$
> $\text{Red}\,\alpha := \beta$

Evidently $\text{Red}\,\alpha \in \mathcal{S}$. Now it is easy to extend the addition to $\mathcal{S}$.

Let $\alpha, \beta \in \mathcal{S}$ then

$$\alpha \oplus \beta = \text{Red}\,(\alpha \cup \beta)$$

$(\mathcal{S}, \oplus, \omega)$ is an idempotent Abelian monoid (semilattice) [3].

A neutral element for the addition is the empty set $\emptyset$ which can also be represented by a special element

$$\omega = \{(\sharp, \emptyset, \sharp, \infty)\}$$

where $\sharp$ represents a "joker"-vertex, which matchs any vertex. We could base the introduction of ω also on the theorem that every semigroup $(S, \circ)$ without a neutral element can be extended to a monoid $(S \cup \{e\}, \circ, e)$ by addition of a new element $e \notin S$ satisfying the relation $\forall x \in S \cup \{e\} : e \circ x = x \circ e = x$.

Now we can also define the *multiplication* of one element sets $\alpha = \{(u, P, v, p)\}$ and $\beta = \{(w, Q, t, q)\}$:

$$\alpha \odot \beta = \begin{cases} \left.\begin{array}{lll} \alpha & q = 0 & \\ \beta & p = 0 & \\ \{(u, P \cup Q, t, p + q)\} & \left.\begin{array}{l} u \neq t,\ P \cap Q = \{v\} \\ u = t,\ P \cap Q = \{v, u\} \end{array}\right\} & v \notin \{u, t\} \end{array}\right\} v = w \\ \omega \qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

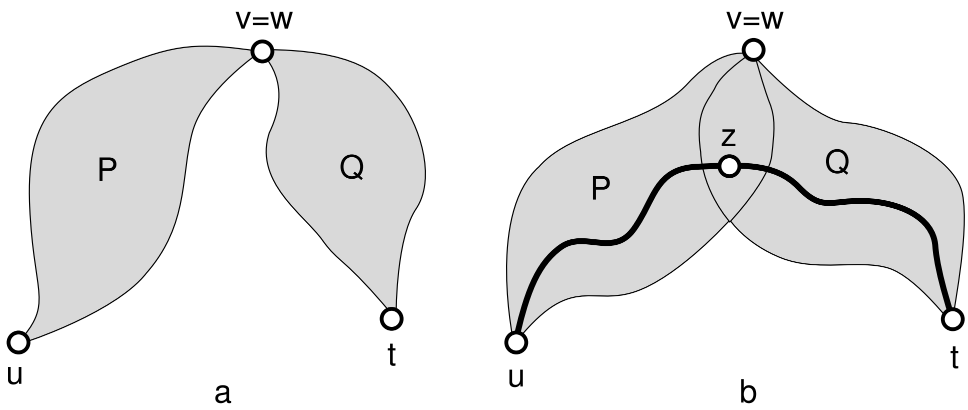


Figure 4: Geosetic multiplication

This definition needs some explanation. The product $\alpha \odot \beta$ is different from ω only if the terminal vertex v of α and the intial vertex w of β coincide (see Figure 4a). Assume that $v \notin \{u, t\}$. If $u \neq t$ and there exists $z \in P \cap Q \setminus \{v\}$ then there also exists a shorter u-t path on $P \cup Q$ (see Figure 4b). Therefore the u-t paths through v are not geodesics. This also covers the case $z = u$ or $z = t$. For example

$$\{(u, \{u, x, v\}, v, 2)\} \odot \{(v, \{v, y, u, t\}, t, 3)\} = \omega$$

Other cases are dealt with similarly.

We extend the multiplication to $\mathcal{S}$ by

$$\alpha \odot \beta = \bigoplus_{a \in \alpha, b \in \beta} \{a\} \odot \{b\}$$

Since for

$$\varepsilon = \{(v, \{v\}, v, 0) : v \in V\}$$

$(\mathcal{S}, \odot, \varepsilon)$ is a monoid and the multiplication $\odot$ distributes over the addition $\oplus$, the structure $(\mathcal{S}, \oplus, \odot, \omega, \varepsilon)$ is an idempotent semiring; and since $\mathcal{S}$ is finite it is also a complete semiring.

The neutral element for addition ω is an absorptive element for multiplication – a *zero*: for $\alpha \in \mathcal{S}$ we have

$$\omega \odot \alpha = \alpha \odot \omega = \omega$$

From the definition of multiplication it follows for $\alpha = \{(u, P, v, p)\},\ p > 0$ that

$$\alpha^2 = \alpha \odot \alpha = \omega$$

Therefore, in this case, $\alpha^k = \omega,\ k \geq 2$ and

$$\alpha^\star = \bigoplus_{k \in \mathbb{N}} \alpha^k = \varepsilon \oplus \alpha$$

In the case of a "diagonal element" $\alpha = \{(v, P, v, p)\}$ (used in Fletcher's algorithm) we have further $\alpha^\star = \varepsilon \oplus \alpha = \varepsilon$.

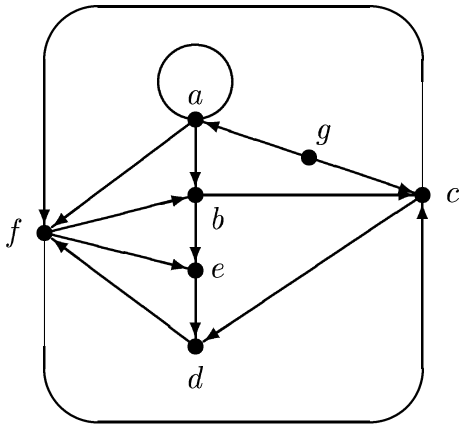


Figure 5: Example graph

	a	b	c	d	e	f	g
a	$(a,1)$	$(ab,1)$	$(abcf,2)$	$(abcdef,3)$	$(abef,2)$	$(af,1)$	$(\emptyset,\infty)$
b	$(\emptyset,\infty)$	$(bcf,3)$	$(bc,1)$	$(bcde,2)$	$(be,1)$	$(bcf,2)$	$(\emptyset,\infty)$
c	$(\emptyset,\infty)$	$(bcf,2)$	$(cf,2)$	$(cd,1)$	$(cef,2)$	$(cf,1)$	$(\emptyset,\infty)$
d	$(\emptyset,\infty)$	$(bdf,2)$	$(cdf,2)$	$(cdef,3)$	$(def,2)$	$(df,1)$	$(\emptyset,\infty)$
e	$(\emptyset,\infty)$	$(bdef,3)$	$(cdef,3)$	$(de,1)$	$(def,3)$	$(def,2)$	$(\emptyset,\infty)$
f	$(\emptyset,\infty)$	$(bf,1)$	$(cf,1)$	$(cdef,2)$	$(ef,1)$	$(cf,2)$	$(\emptyset,\infty)$
g	$(ag,1)$	$(abg,2)$	$(cg,1)$	$(cdg,2)$	$(abcefg,3)$	$(acfg,2)$	$(\emptyset,\infty)$

Table 6: Geosetic closure

4.2 Algorithm and example

This semiring can be used to compute Cayley's table of Boyle's operation for a given network by computing a strict closure of a relation matrix over it.

First we transform relation R to a matrix $\mathbf{R} = [(P,p)_{u,v}]$ which has for the entries pairs defined by

$$(P,p)_{u,v} = \begin{cases} (\{u,v\},1) & (u,v) \in R \\ (\emptyset,\infty) & (u,v) \notin R \end{cases}$$

where P is the set of vertices on the shortest u-v paths and p is their length.

Since the addition is idempotent we can apply the Fletcher's algorithm [14] in place. The term $(c_{k-1}[k,k])^\star = \varepsilon$ and therefore it can be omitted.

Representing the relation matrix **R** (and its strict closure) by two matrices: the matrix **S** of sets of vertices on the shortest paths, and the shortest paths lengths matrix $\mathbf{D}$, we obtain the following adapted version of Fletcher's algorithm:

for $k := 1$ **to** n **do begin**
    for $i := 1$ **to** n **do for** $j := 1$ **to** n **do begin**
        $piq := s[i,k] * s[k,j]$;
        if $(piq - [i]) = [k]$ **then begin**
            $puq := s[i,k] + s[k,j]$;
            $dst := d[i,k] + d[k,j]$;
            if $d[i,j] = dst$ **then** $s[i,j] := s[i,j] + puq$
            else if $d[i,j] > dst$ **then**
                begin $s[i,j] := puq; d[i,j] := dst$ **end**;
        end;
    end;
end;

Note, for sets in pascal $+$ denotes the union $\cup$, and $*$ denotes the intersection $\cap$.

Using this algorithm we obtained for a graph represented in Figure 5 a strict geosetic closure presented in Table 6.

5 Conclusion

In the paper four semirings for solving social networks analysis problems were constructed. For balance and cluster semirings Fletcher's algorithm for computing the closure of a network matrix can be applied directly; and for geodetic and geosetic semirings we have adapted versions of Fletcher's algorithm. The algorithm for geodetic closure is more efficient than the usual Harary, Norman, Cartwright algorithm [19]. We do not know any other algorithm for computing Cayley's table of Boyle's operation (geosetic closure).

The embedding technique used in the construction of a geosetic semiring can be used also as a general framework for path problems in valued graphs.

References

[1] A. V. Aho, J. E. Hopcroft, J. D. Ullman, *The Design and Analysis of Computer Algorithms.* Addison-Wesley, Reading, Massachusetts (1976).

[2] Batagelj V., "Signed Graphs and Closed Semirings." Network Seminar, Department of Sociology, University of Pittsburgh, December 1990, manuscript.

[3] Batagelj V., "Use of Mathematica in verifying the algebraic structure axioms – semirings." In preparation.

[4] Boyle R.: "Algebraic systems for normal and hierarchical sociograms." *Sociometry* **32** (1969), 91-111.

[5] R.E. Burkard, R.A. Cuninghame-Greene, U. Zimmermann, eds., *Algebraic and Combinatorial Methods in Operations Research.* Annals of Discrete Mathematics **19** (1984).

[6] B. Carré, *Graphs and Networks.* Clarendon, Oxford (1979).

[7] D. Cartwright, F. Harary, *On the coloring of signed graphs.* Elemente der Mathematik **23**(1968), pp. 85-89.

[8] D. Cartwright, F. Harary, *Balance and Clusterability: An Overview,* Holland P.W., Leinhardt S. (eds.) Perspectives on Social Network Research, Academic Press, New York (1979), pp. 25-50.

[9] G. Chartrand, *Introductory Graph Theory.* Dover, New York (1985).

[10] J. A. Davis, "Clustering and structural balance in graphs." *Human Relations,* **20**(1967), pp. 181-187.

[11] P. Doreian, *Mathematics and the study of social relations.* Weidenfeld and Nicolson, London (1970).

[12] P. Doreian, "On the connectivity of social networks." *Journal of Mathematical Sociology* **3**(1974), 248.

[13] Everett M. G.: "A graph theoretic blocking procedure for social networks". *Social networks* **4** (1982), 147-167.

[14] J. G. Fletcher, "A more general algorithm for computing closed semiring costs between vertices of a directed graph," *CACM* (1980), pp. 350-351.

[15] L. C. Freeman, "Centrality in social networks: Conceptual clarification." *Social Networks,* **1**(1978/79), 215-239.

[16] R. V. Gould, "Measures of betweenness in non-symmetric networks. Social Networks, **9**(1987), 277-282.

[17] P. Hage P., F. Harary F., *Structural models in anthropology.* Cambridge University Press, Cambridge (1983).

[18] F. Harary, "On the notion of balance in signed graphs". *Michigan Mathematical Journal* (1953) 2, pp. 143-146.

[19] F. Harary, R. Z. Norman, D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs.* John Wiley & Sons, New York (1965).

[20] J. E. Riley, "An application of graph theory to social psychology." *The Many Facets of Graph Theory*, Lecture Notes in Mathematics, 110, (G. Chartrand and S. F. Kapoor, eds.). Springer-Verlag, Berlin (1969), pp. 275-280.

[21] F. S. Roberts, *Discrete Mathematical Models.* Prentice-Hall, Englewood Cliffs, New Jersey (1976).

[22] U. Zimmermann, *Linear and Combinatorial Optimization in Ordered Algebraic Structures.* Annals of Discrete Mathematics **10** (1981).