

A Privacy-Aware Protocol for Sociometric Questionnaires

Marián Novotný

Institute of Computer Science
P.J. Šafárik University, Faculty of Science
Košice, Slovakia

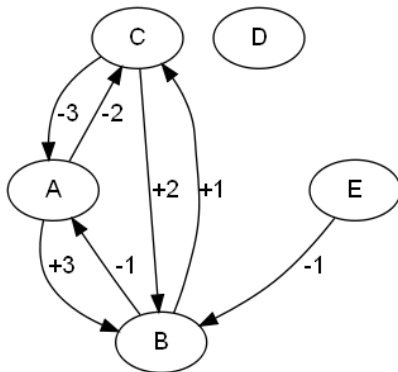
MEMICS 2009 - 5th Doctoral Workshop on Mathematical
and Engineering Methods in Computer Science

Introduction to Sociometry

- quantitative method for measuring **social relationships** (Jacob L. Moreno)
- can be used for management of a school class by a teacher or in a team-building
- is based on **choices** of individuals
 - responders are asked to choose one or more persons from the group according to specific **criteria**
 - choices of responders are collected by a **questionnaire** from responders
- relations between individuals can be represented by a **sociogram**

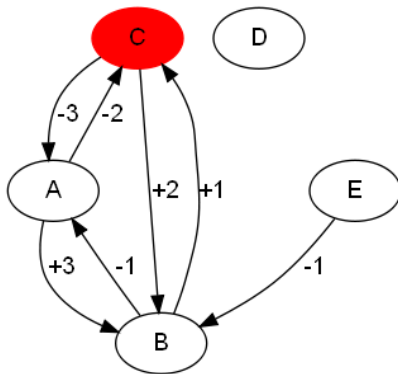
Representation of a Sociogram by Graph Theory

- **weighted digraph** $G = (V, E)$, $E \subseteq V \times V$,
 - **nodes** from V represent individuals from the social group
 - social link is represented as a **weighted arc** from E
- the weight function $w : E \rightarrow \{-s, \dots, -1, 1, \dots, s\}$
 - expresses **rates** of social links



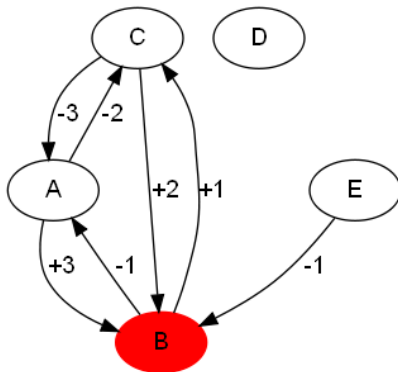
Node Characteristics – Indegrees

- **positive indegree** $deg^{In^+}(C) = 1$
- **negative indegree** $deg^{In^-}(C) = 1$
- **indegree** $deg^{In}(C) = deg^{In^+}(C) + deg^{In^-}(C) = 2$



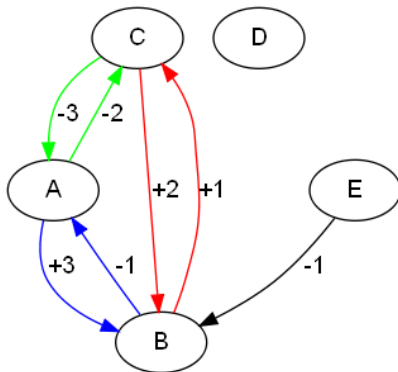
Node Characteristics – Weighted Indegrees

- positive weighted indegree $In^+(B) = 5$
- negative weighted indegree $In^-(B) = -1$



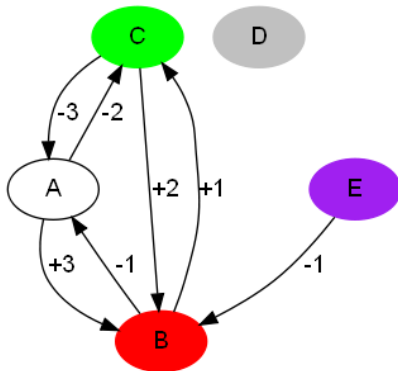
Mutual Choices

- positive mutual choice
- negative mutual choice
- combined mutual choice



Individual Phenomena

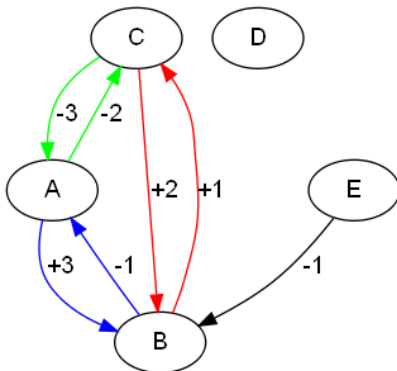
- **positive social status** of B $\frac{In^+(B)}{|V|-1} = \frac{5}{4}$
- **Star** B – node with the maximal positive weighted indegree
- **Outsider** C – node with the minimal negative weighted indegree
- **Ghost** D – node with zero indegree and outdegree
- **Isolate** E – node with zero positive indegree, is not a ghost



Collective Phenomena

- the set of **positive** M^+ , **negative** M^- , **combined** M^\pm mutual choices
- positive coherence** of a group G is defined as

$$\text{coh}^+(G) = \frac{|M^+|}{\binom{|V|}{2}} = \frac{1}{\binom{5}{2}} = \frac{1}{10}$$



Security Requirements for the Scheme

- **Eligibility** – only responders from the group are eligible to correctly fill in the questionnaire.
- **Privacy** – choices of a responder must not identify the responder and any traceability between the responder and his choices must be removed.
- **Verifiability** –
 - a responder should be able to **individually** verify whether his choices were correctly recorded and accounted
 - participants can **universally** verify that in the evaluation process only valid choices of eligible responders were recorded and the counting process was accurate
- **Accuracy** – the scheme must be error-free. The final computation of sociometric indices must correspond with all choices of all responders.

The Homomorphic Public-Key System

- used for encryption of responders choices
- **semantically** secure, additively **homomorphic**, allows us once to use multiplication
- **robust threshold** version (t, a)
 - the **private key** is shared among a authorities
 - A ciphertext can be decrypted when at least $t + 1$ shareholders cooperate
 - the process of decryption is **universally verifiable** and does not reveal the secret key

Homomorphic Properties of the Public-Key System

- given ciphertexts $C_1 = E_{PK}(m_1)$, $C_2 = E_{PK}(m_2)$, **anyone** can create
 - $E_{PK}(m_1 + m_2)$ by computing the **product**
 $C_1 \cdot C_2 = E_{PK}(m_1 + m_2)$
 - $E_{PK}(m_1 \cdot m_2)$ by computing the **bilinear map**
 $C_1 \star C_2 = E_{PK}(m_1 \cdot m_2)$
 - $E_{PK}(z \cdot m_1)$ by computing the **exponentiation**
 $C_1^z = E_{PK}(z \cdot m_1)$

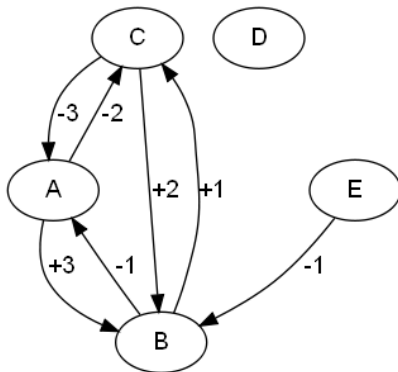
The Proposed Scheme – Registration, Key Generation

- a **trusted dealer** first generates the public key Pk and the private key Sk , shares the private keys between a authorities and then deletes the private key
- the questioner **creates** a questionnaire and registers it by the **collector**. Questionnaire contains obligatory properties
 - the **list of responders** with their unique identification
 - **sociometric indices** which have to be computed
 - the **deadline** for filling and the sociometric parameters
- A responder
 - **selects** his choices
 - **submits** signed selections encrypted under the key Pk to the collector
- Collector
 - **collects** submissions of responders and **checks** signatures
 - **leads** the computations and publishes submissions and results

The Proposed Scheme – Representation of Arcs

- to **represent** a weighted arc from the node R_i to node R_j we use $s + 2$ bits $b_{ij}^+, b_{ij}^-, b_{ij}^{w_1}, \dots, b_{ij}^{w_s}$

| | |
|-------------------|-------|
| $A \rightarrow B$ | 10001 |
| $A \rightarrow C$ | 01010 |
| no arc | 00100 |



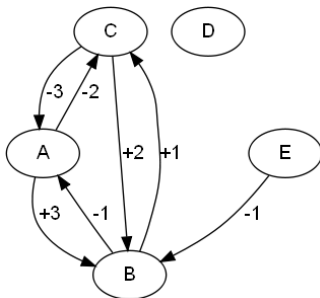
The Proposed Scheme – Verification of Submissions

- to represent a weighted arc from the node R_i to node R_j we use $s + 2$ bits $b_{ij}^+, b_{ij}^-, b_{ij}^{w_1}, \dots, b_{ij}^{w_s}$
- we need to **verify**, that
 - $b_{ij}^\diamond \in \{0, 1\} \equiv b_{ij}^\diamond \cdot (b_{ij}^\diamond - 1) = 0$
 - $b_{ij}^+ \cdot b_{ij}^- = 0$
 - $\sum_{k=1}^s b_{ij}^{w_k} = 1 \equiv \sum_{k=1}^s b_{ij}^{w_k} - 1 = 0$
- We use the homomorphic properties for preparing ciphertexts of **equations**
- The equations can be checked by shareholders by cooperatively-made **decryptions**
- to save on computation, we check at once a **batch of equations**

The Proposed Scheme – Encrypted Sociogram

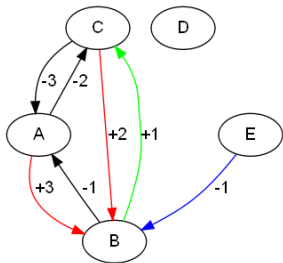
- $$c_{ij}^w = \prod_{k=1}^S (c_{ij}^{w_k})^k == E_{PK}(\sum_{k=1}^S k \cdot b_{ij}^{w_k}) = E_{PK}(|w_{ij}|)$$

| | A | B | C | D | E |
|---|--------------------|--------------------|--------------------|--------------------|--------------------|
| A | — | $E(1), E(0), E(3)$ | $E(0), E(1), E(2)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| B | $E(0), E(1), E(1)$ | — | $E(1), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| C | $E(0), E(1), E(3)$ | $E(1), E(0), E(2)$ | — | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| D | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — | $E(0), E(0), E(1)$ |
| E | $E(0), E(0), E(1)$ | $E(0), E(1), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — |



Computation of Characteristics of Nodes

| | A | B | C | D | E |
|---|--------------------|--------------------|--------------------|--------------------|--------------------|
| A | — | $E(1), E(0), E(3)$ | $E(0), E(1), E(2)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| B | $E(0), E(1), E(1)$ | — | $E(1), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| C | $E(0), E(1), E(3)$ | $E(1), E(0), E(2)$ | — | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| D | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — | $E(0), E(0), E(1)$ |
| E | $E(0), E(0), E(1)$ | $E(0), E(1), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — |



$$E(\text{deg}^{\text{In}^+}(B)) = E(1) \cdot E(1) \cdot E(0) \cdot E(0) = E(2)$$

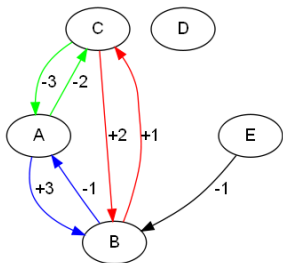
$$E(\text{deg}^{\text{In}^-}(B)) = E(0) \cdot E(0) \cdot E(0) \cdot E(1) = E(1)$$

$$E(\text{deg}^{\text{Out}^+}(B)) = E(0) \cdot E(1) \cdot E(0) \cdot E(0) = E(1)$$

$$E(\text{In}^+(B)) = (E(1) \star E(3)) \cdot (E(1) \star E(2)) \cdot (E(0) \star E(1)) \cdot (E(0) \star E(1)) = E(3 \cdot 1 + 2 \cdot 1 + 0 \cdot 1 + 0 \cdot 1) = E(5)$$

Computation of the Mutual Choices

| | A | B | C | D | E |
|---|--------------------|--------------------|--------------------|--------------------|--------------------|
| A | — | $E(1), E(0), E(3)$ | $E(0), E(1), E(2)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| B | $E(0), E(1), E(1)$ | — | $E(1), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| C | $E(0), E(1), E(3)$ | $E(1), E(0), E(2)$ | — | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ |
| D | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — | $E(0), E(0), E(1)$ |
| E | $E(0), E(0), E(1)$ | $E(0), E(1), E(1)$ | $E(0), E(0), E(1)$ | $E(0), E(0), E(1)$ | — |



$$\prod_{i=1}^N \prod_{j>i} c_{ij}^+ * c_{ji}^+ = \prod_{i=1}^N \prod_{j \in J_i} E_{Pk}(b_{ij}^+ b_{ji}^+) =$$

$$\prod_{i=1}^N E_{Pk}(\sum_{j \in J_i} b_{ij}^+ b_{ji}^+) = E_{Pk}(\sum_{i=1}^N \sum_{j \in J_i} b_{ij}^+ b_{ji}^+) =$$

$$E_{Pk}(|M^+|)$$

Computations in the Bit-Representation of a Sociogram

| | | |
|--|---|--|
| $\text{deg}^{In^+}(R_i) = \sum_{j \in J_i} b_{ji}^+$ $\text{deg}^{Out^-}(R_i) = \sum_{j \in J_i} b_{ij}^-$ $\text{In}^+(R_i) = \sum_{j \in J_i} b_{ji}^+ \cdot w_{ji} $ | $\text{deg}^{In^-}(R_i) = \sum_{j \in J_i} b_{ji}^-$ $\text{deg}^{In}(R_i) = \sum_{j \in J_i} b_{ji}^+ + b_{ji}^-$ $\text{In}^-(R_i) = - \sum_{j \in J_i} b_{ji}^- \cdot w_{ji} $ | $\text{deg}^{Out^+}(R_i) = \sum_{j \in J_i} b_{ij}^+$ $\text{deg}^{Out}(R_i) = \sum_{j \in J_i} b_{ij}^+ + b_{ij}^-$ |
| $ M^+ = \sum_{i=1}^N \sum_{j>i} b_{ij}^+ \cdot b_{ji}^+$ $ M^- = \sum_{i=1}^N \sum_{j>i} b_{ij}^- \cdot b_{ji}^-$ | $ M^\pm = \sum_{i=1}^N \sum_{j>i} (b_{ij}^- \cdot b_{ji}^+) + (b_{ij}^+ \cdot b_{ji}^-)$ $ M = \sum_{i=1}^N \sum_{j>i} (b_{ij}^- \cdot b_{ji}^+) + (b_{ij}^+ \cdot b_{ji}^-) + (b_{ij}^+ \cdot b_{ji}^+) + (b_{ij}^- \cdot b_{ji}^-)$ | |

- we **derive** encrypted values of indices **only** using **homomorphic** properties of the encryption system

Conclusions

- we proposed a representation of a sociogram by a **weighted digraph**
- we proposed the threshold version of the public key system
- we designed the **protocol** for anonymous sociometric questionnaires
 - based on **additively homomorphic** public key cryptosystem, which allows us **once** to use **multiplication**
 - to prepare the submission costs $(N - 1)(s + 2)$ encryptions and one digital signature
 - fulfils desired security requirements
 - eligibility – digital signatures, verification of submissions
 - privacy – semantically secure cryptosystem
 - verifiability: individually –checking published list of submissions, universally by verification of signatures, computations and decryptions
 - accuracy - we derived encrypted values of indices **only** using **homomorphic** properties

Thank you for your attention