University at Buffalo
The State University of New York

Department of Industrial and Systems Engineering
School of Engineering and Applied Sciences

Bayesian Evidence Cascades and Seed-Initiated Marketing Campaigns in Social Networks

Mohammadreza Samadi
PhD Candidate,
Department of Industrial and Systems Engineering
University at Buffalo, SUNY

Alexander Nikolaev
Assistant Professor,
Department of Industrial and Systems Engineering
University at Buffalo, SUNY

Rakesh Nagi
Donald Biggar Willett Professor and Head
Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

May, 2015

Outline

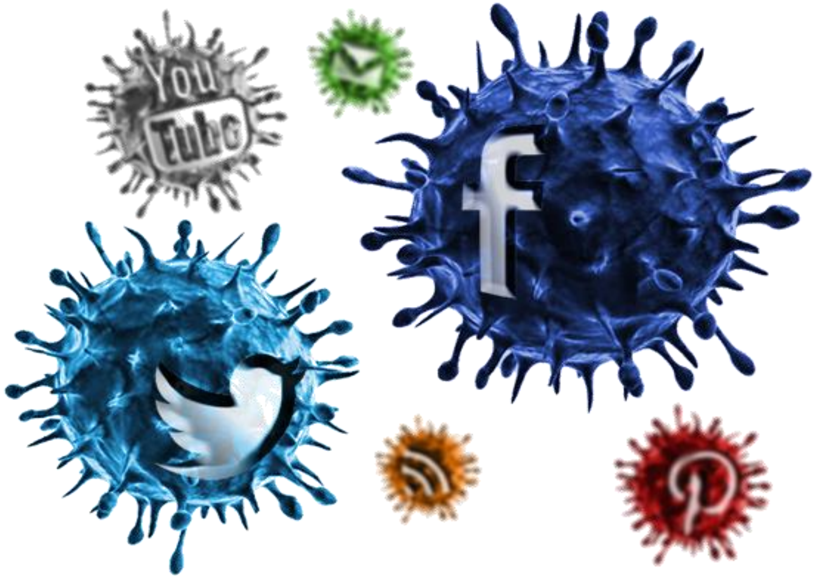
- Introduction and Motivation (Influence Maximization)
- Parallel Cascade (PC) Diffusion Model
- Case Studies
- Mathematical Formulation
- Lagrangian Relaxation Heuristic
- Seed Selection Scheduling Problem
- Appropriate Time Horizon for Cascades
- Conclusion

Word-of-Mouth and Viral Marketing

- Knowledge transfer between individuals affects their purchasing/voting decisions.
- 69% of consumers consult friends and family before purchasing home electronics.

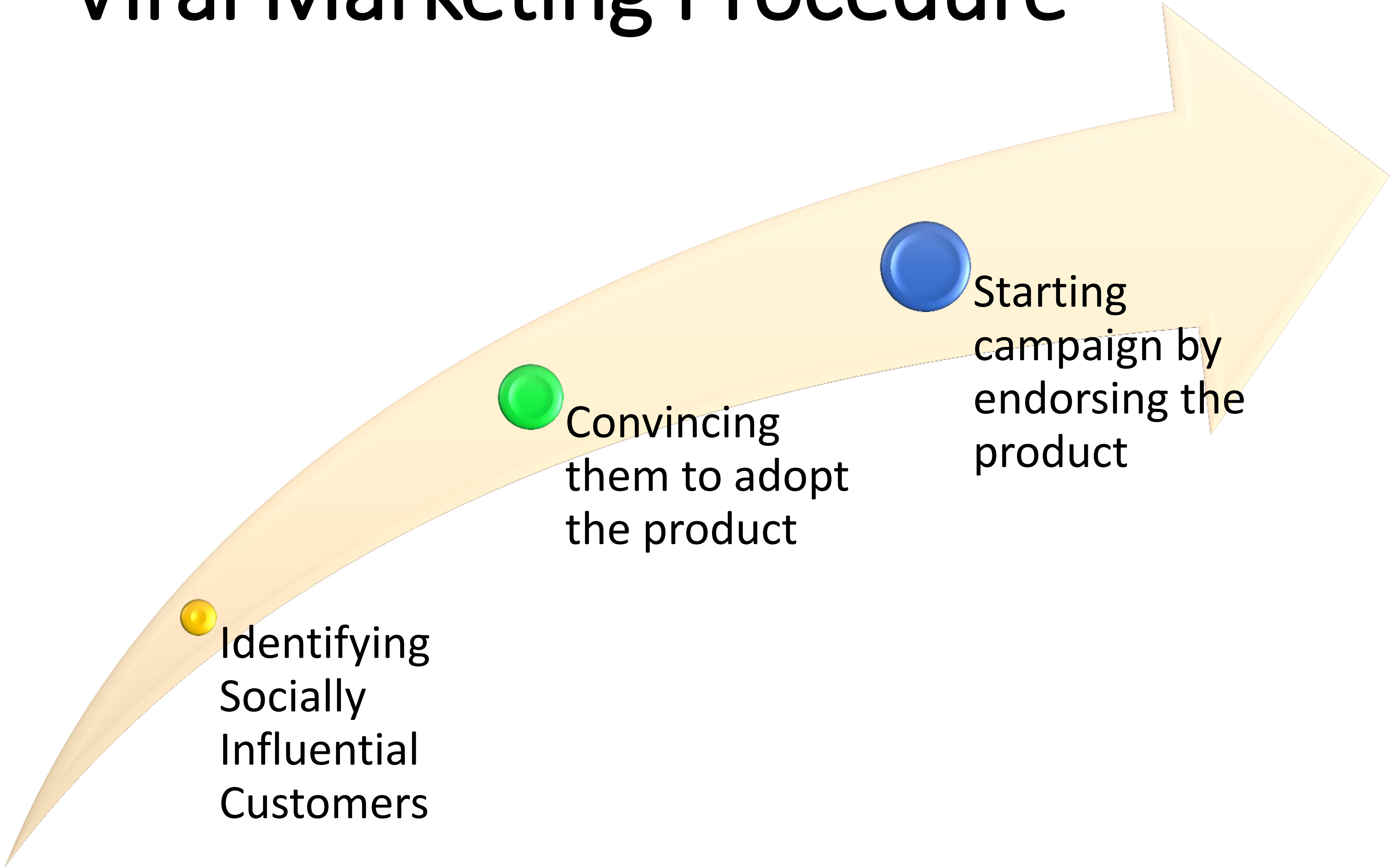
Word-of-Mouth and Viral Marketing

- The main element for creating and controlling cascades is the set of early adopters.
- Early adopters, called **seeds**, are motivated by offering discounted/free products to start the cascades in the network.



- Marketing/campaign budget determines how many seeds can be selected.
- Online social networks provide a good opportunity for network marketing.

Viral Marketing Procedure



Identifying Socially Influential Customers

Convincing them to adopt the product

Starting campaign by endorsing the product

A pool of adopted customers

Question

How to select the seeds?!

Influence Maximization Problem answers this question

Diffusion Model

- Defines how the influence is propagated in the network.
- Simulation Process!

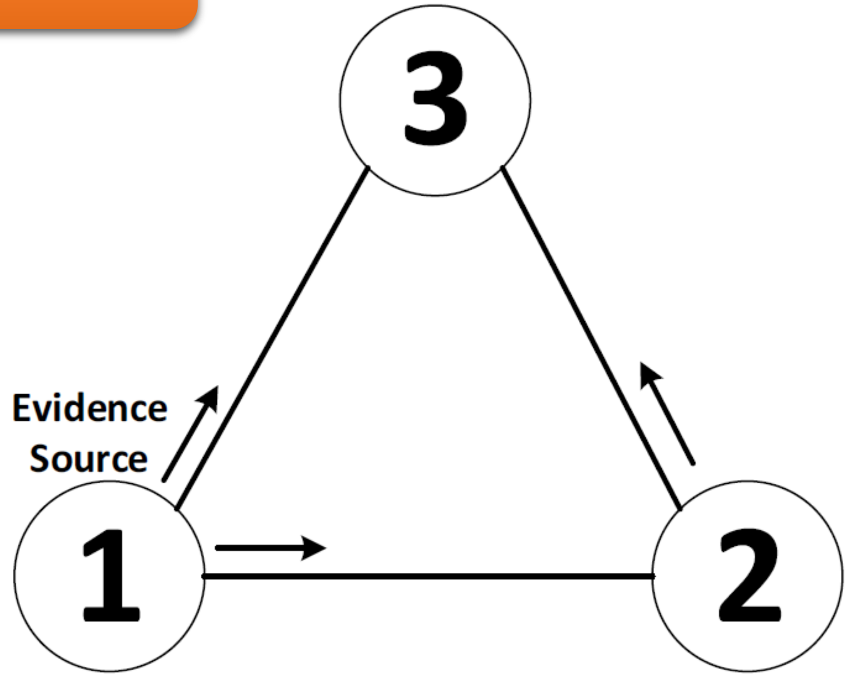
Optimization Model

- Designed based on the diffusion model.
- Finds the optimal set of seeds to maximize the spread of influence.

(Subjective) Bayesian Evidence Spread

Null hypothesis: a new phone service is reliable.

- Node 1 experiences no dropped calls in a month.
- Node 1 presents the impression to nodes 2 and 3.
- Both nodes 2 and 3 update their beliefs.
- Node 2 transfers information to node 3 without providing the source of information.
- Node 3 treats the new information as if it provides new evidence supporting the hypothesis



Parallel Cascade Diffusion Model

- Modeling the parallel duplication of influence in the typology of flow processes on social networks

Social network: directed finite graph *G*=(*V*,*E*)

Set of positive seeds must be determined.

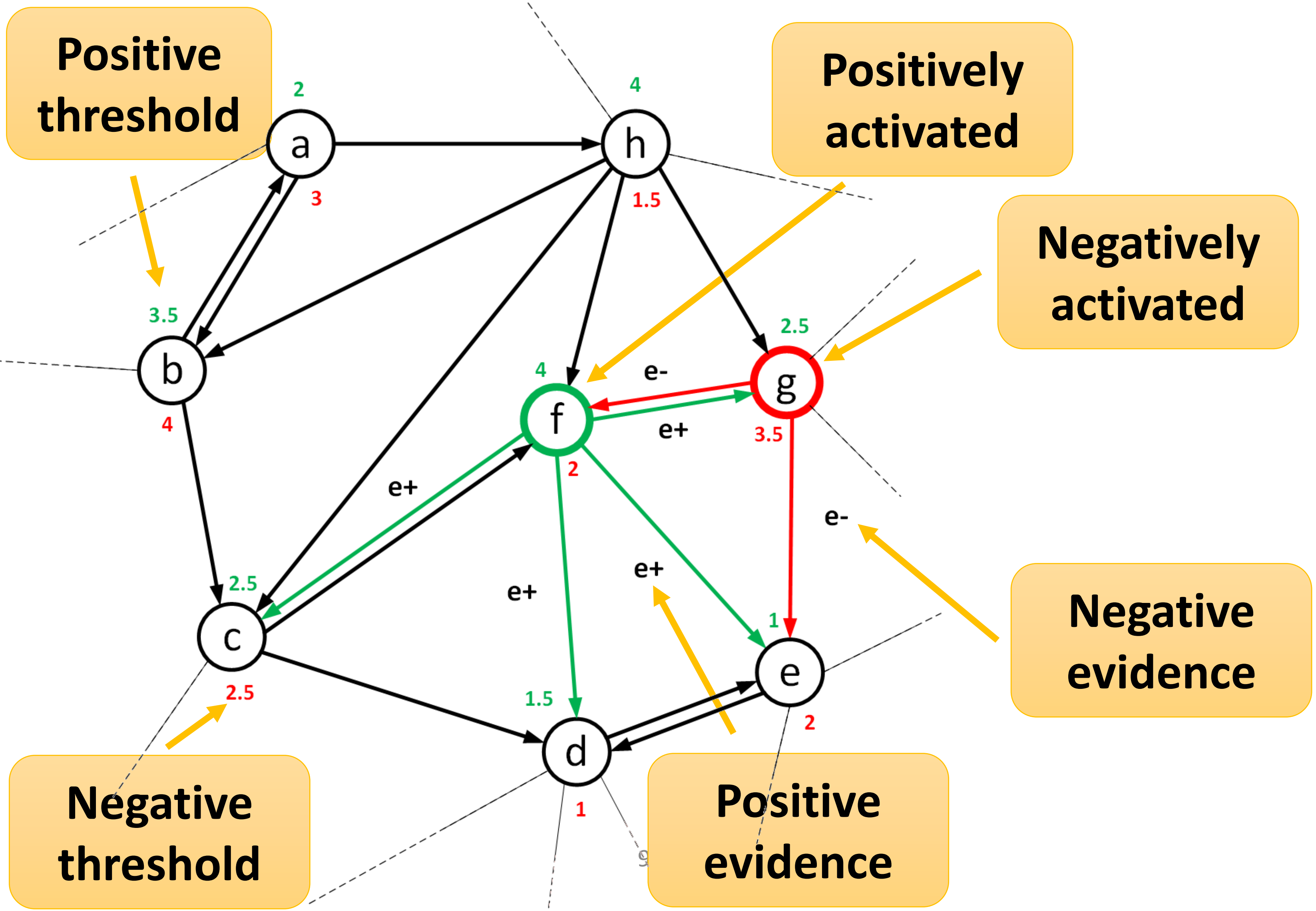
Set of negative seeds is given (competitor's agents in the market)

An attached memory to each node to collect all evidence (activation level).

positive and a negative threshold values

A node is positively (negatively) activated when its activation level passes the positive (negative) threshold.

Parallel Cascade Diffusion Model



Positive threshold

Positively activated

Negatively activated

Negative evidence

Negative threshold

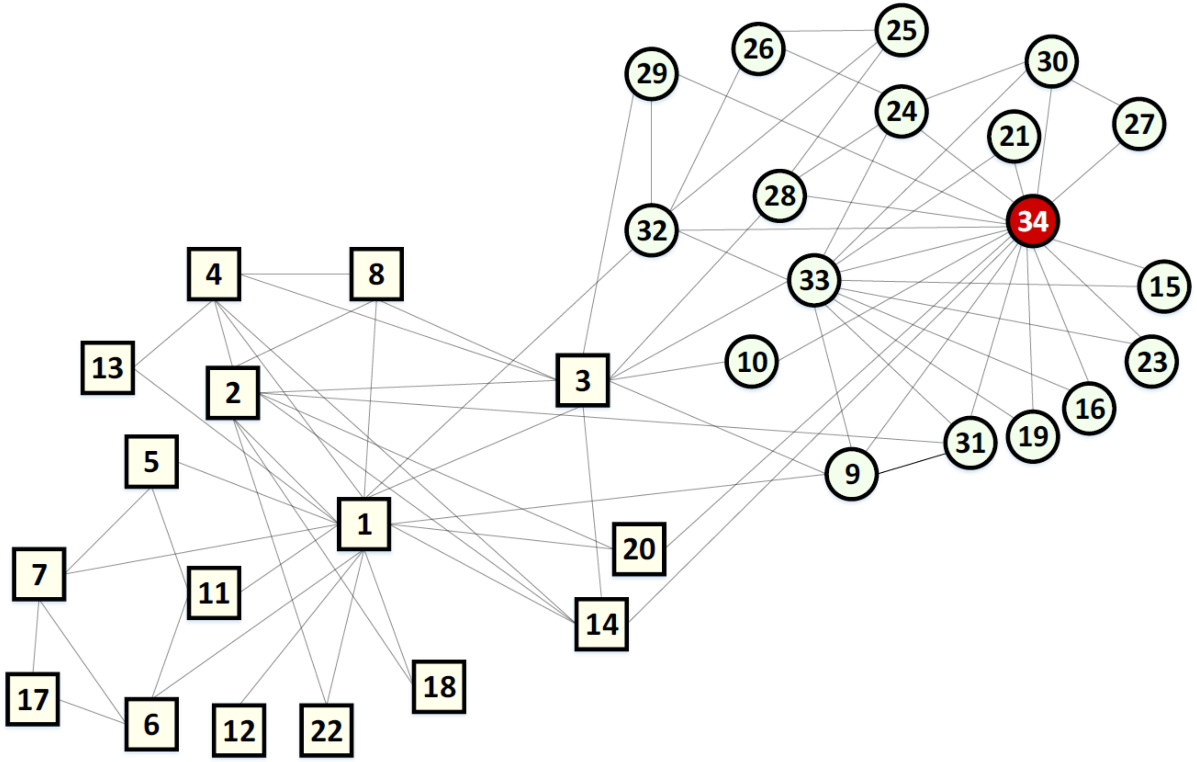
Positive evidence

- The value of the evidence calculated using Bayesian Inference logic.
- Evidence value decreases by rate α.
- A node forgets the previously perceived evidences by rate β.
- Updating the activation level at the end of each time period.
- A node can lose its activation or get the opposite activation.
- The diffusion process is terminated after a predefined number of iterations

Case Study 1: The optimal strategic positioning of the seeds

- Zachary's Karate club members, not yet exposed to a new emerging vitamin supplement product.

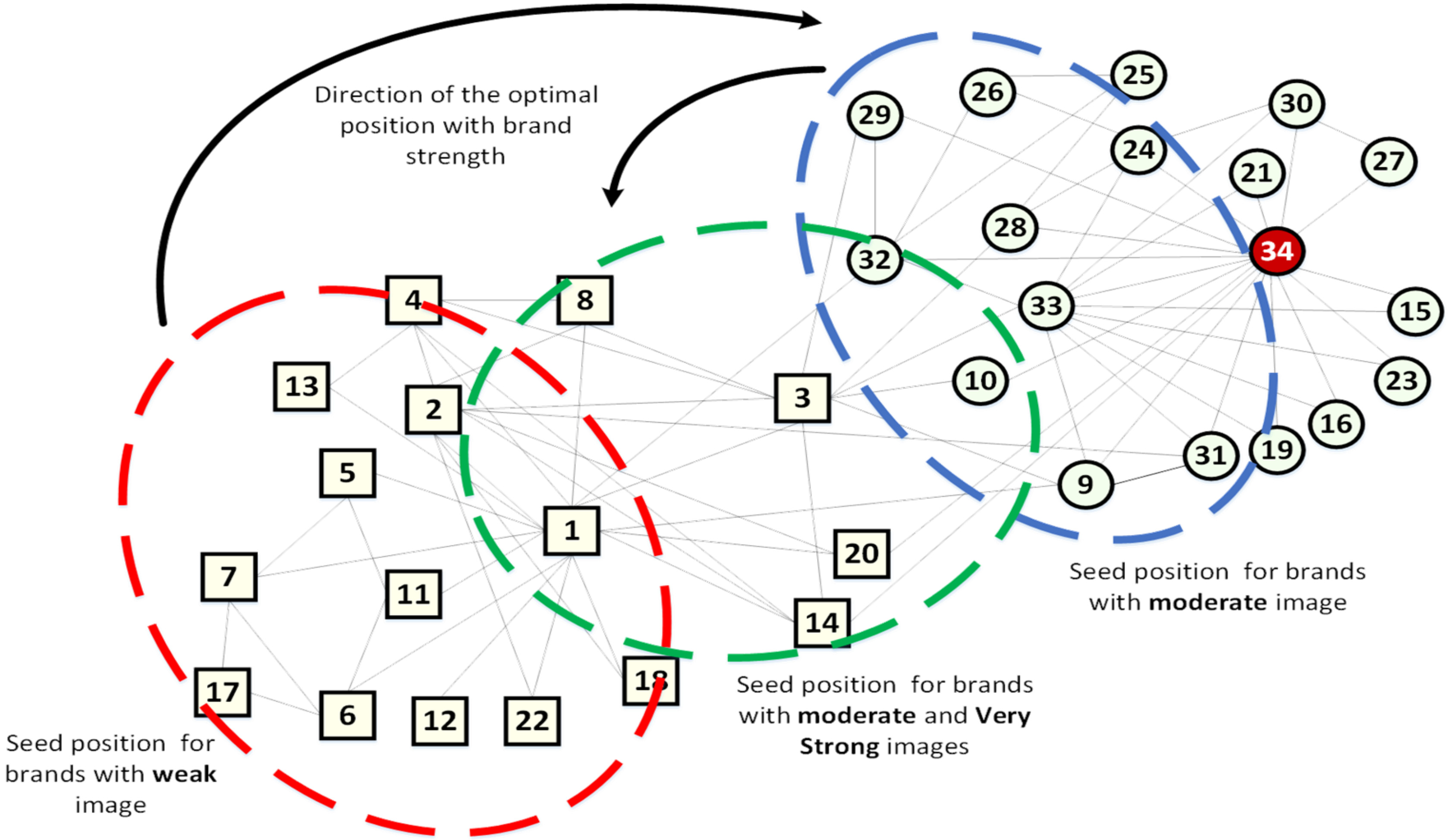
- The decision-maker (producer) seeks to motivate people to use the new product and influence each other.



- The decision-maker plans to offer the product at a discounted price to two members.

Case Study 1: The optimal strategic positioning of the seeds

- The relative strength of brands determines the strategic position of the seeds.



Direction of the optimal position with brand strength

Seed position for brands with **weak** image

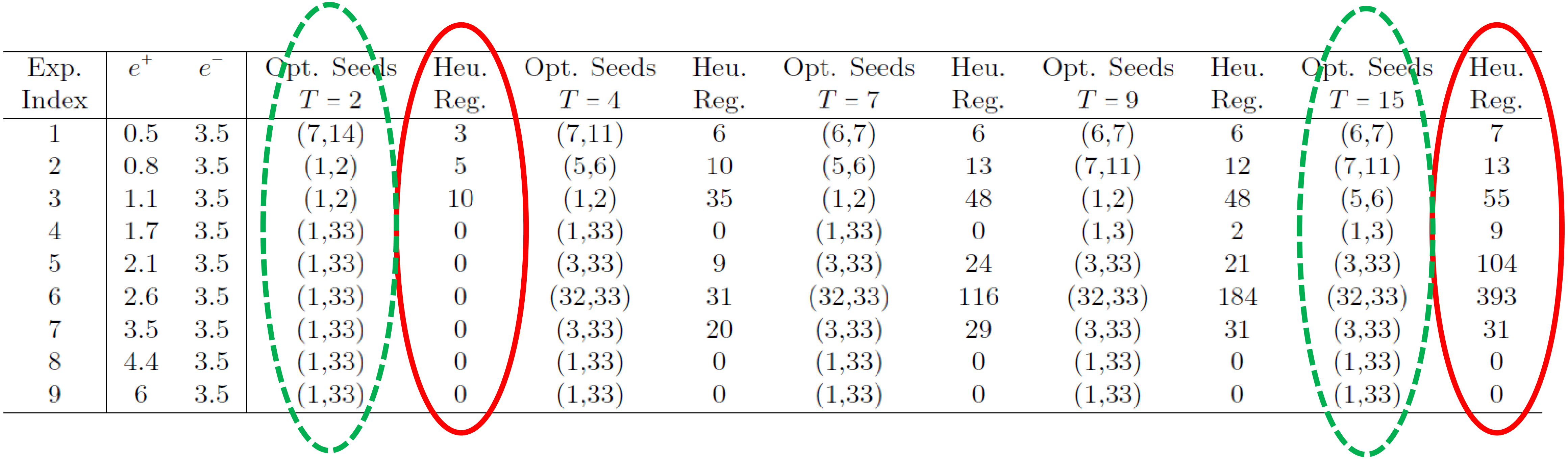
Seed position for brands with **moderate** and **Very Strong** images

Seed position for brands with **moderate** image

Exp. Index	e^+	e^-	Opt. Seeds
1	0.5	3.5	(6,7)
2	0.8	3.5	(5,6)
3	1.1	3.5	(1,2)
4	1.7	3.5	(1,33)
5	2.1	3.5	(3,33)
6	2.6	3.5	(32,33)
7	3.5	3.5	(3,33)
8	4.4	3.5	(1,33)
9	6	3.5	(1,33)

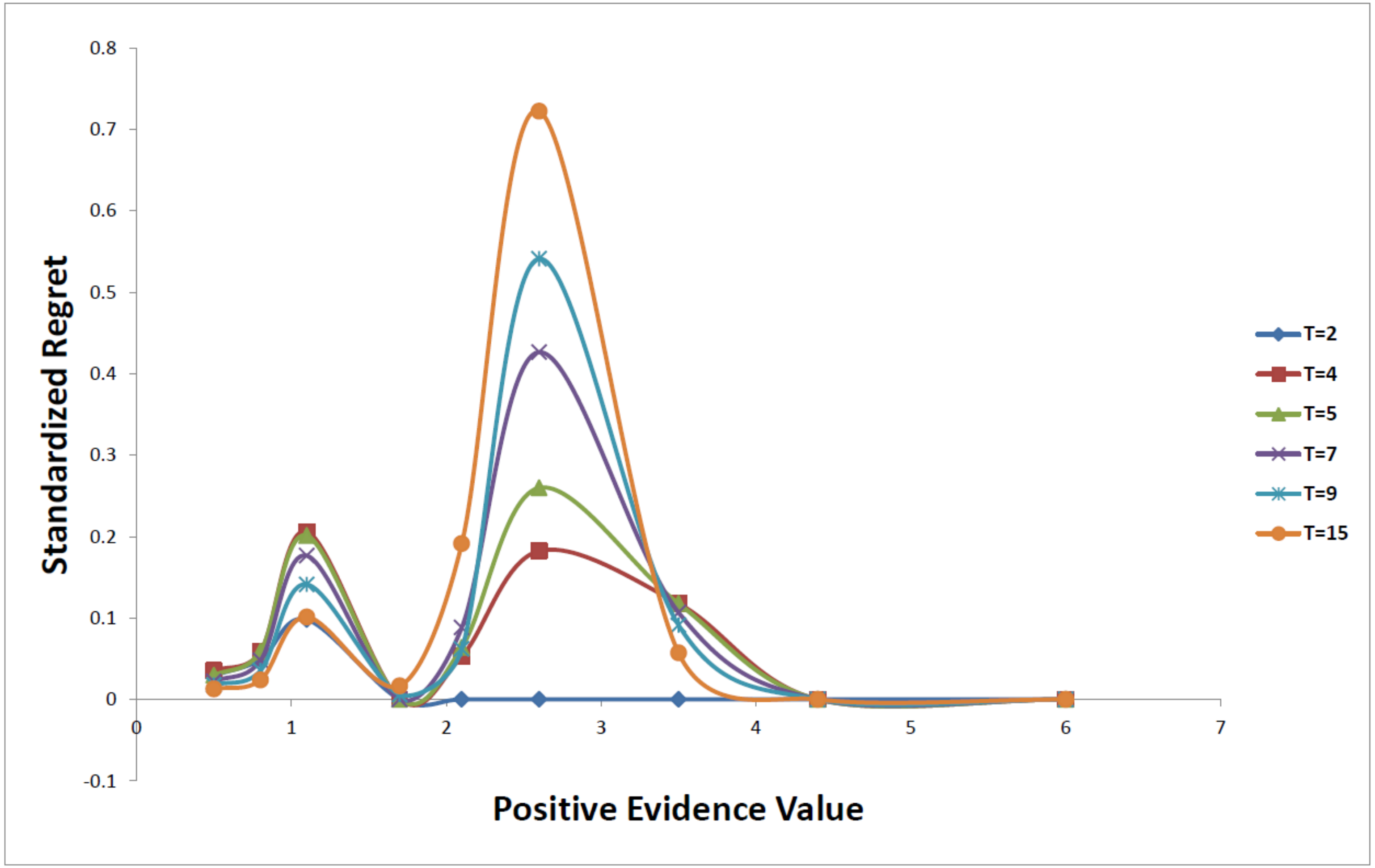
Time horizon and Optimal position of positive seeds

- Increasing time horizon creates a larger regret for the maximum degree heuristic strategy (1,33).



Exp. Index	e^+	e^-	Opt. Seeds $T = 2$	Heu. Reg.	Opt. Seeds $T = 4$	Heu. Reg.	Opt. Seeds $T = 7$	Heu. Reg.	Opt. Seeds $T = 9$	Heu. Reg.	Opt. Seeds $T = 15$	Heu. Reg.
1	0.5	3.5	(7,14)	3	(7,11)	6	(6,7)	6	(6,7)	6	(6,7)	7
2	0.8	3.5	(1,2)	5	(5,6)	10	(5,6)	13	(7,11)	12	(7,11)	13
3	1.1	3.5	(1,2)	10	(1,2)	35	(1,2)	48	(1,2)	48	(5,6)	55
4	1.7	3.5	(1,33)	0	(1,33)	0	(1,33)	0	(1,3)	2	(1,3)	9
5	2.1	3.5	(1,33)	0	(3,33)	9	(3,33)	24	(3,33)	21	(3,33)	104
6	2.6	3.5	(1,33)	0	(32,33)	31	(32,33)	116	(32,33)	184	(32,33)	393
7	3.5	3.5	(1,33)	0	(3,33)	20	(3,33)	29	(3,33)	31	(3,33)	31
8	4.4	3.5	(1,33)	0	(1,33)	0	(1,33)	0	(1,33)	0	(1,33)	0
9	6	3.5	(1,33)	0	(1,33)	0	(1,33)	0	(1,33)	0	(1,33)	0

Regret of heuristic approach



- The regret needs a standardization.
- The standard regret is larger for longer time horizons
- Increasing the time horizon makes the heuristic strategies less attractive.

Mixed-Integer Program

- The optimization problem just decides about the position of the positive seeds so that maximize the spread of positive seeds and minimize the spread of negative seeds.

- The objective function cares about both activation status and time of activation.

Number of nodes

Number of time periods

$$(P) \qquad \max Z = \sum_{i=1}^{N} \sum_{t=0}^{T} (X_{it} - Y_{it})$$

Negative activation variable

Positive activation variable

Mathematical Model

Activation Constraints

$$Y_{it} \geq ((K_{it} - L_{it}) - \theta_i^-)/M \qquad i = 1, 2, ...|N|, \quad t = 0, 1, ..., T,$$

$$1 - X_{it} \geq (\theta_i^+ - (L_{it} - K_{it}))/M \qquad i = 1, 2, ...|N|, \quad t = 0, 1, ..., T,$$

$$X_{it} + Y_{it} \leq 1 \qquad i = 1, 2, ...|N|, \quad t = 0, 1, ..., T,$$

Evidence Level Updates

$$L_{it} = \beta_1 L_{it-1} + \sum_{(j,i) \in A} E^+_{jt-1} \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

$$K_{it} = \beta_2 K_{it-1} + \sum_{(j,i) \in A} E^-_{jt-1} \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

$$L_{i0} = X_{i0}(\theta_i^+) \qquad i = 1, 2, ...|N|,$$

$$K_{i0} = Y_{i0}(\theta_i^- + \epsilon) \qquad i = 1, 2, ...|N|,$$

Mathematical Model

$$E^+_{it} \le (\alpha_1 E^+_{it-1}) + (1 - X_{it-1})e^+ \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

Evidence Transferred Value Update

$$E^+_{it} \le e^+(X_{it}) \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

$$E^-_{it} \ge (\alpha_2 E^-_{it-1}) + (Y_{it} - Y_{it-1})e^- \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

$$E^-_{it} \le e^-(Y_{it}) \qquad i = 1, 2, ...|N|, \quad t = 1, 2, ..., T,$$

Initialization

$$Y_{i0} = S^-_i \qquad i = 1, 2, ...|N|,$$

$$E^+_{i0} = X_{i0}e^+ \qquad i = 1, 2, ...|N|,$$

$$E^-_{i0} = Y_{i0}e^- \qquad i = 1, 2, ...|N|,$$

Seed Set Size Constraint

$$\sum_{i=1}^{|N|} X_{i0} \le |S^+|,$$

$$0 \le L_{it}, K_{it}, E^+_{it}, E^-_{it} \qquad i = 1, 2, ...|N|, \quad t = 0, 1, ..., T,$$

$$Y_{it}, X_{it} \in \{0, 1\} \qquad i = 1, 2, ...|N|, \quad t = 0, 1, ..., T.$$

A Guaranteed-Performance Lagrangian Relaxation Heuristic

- Relaxes one of the constraint sets and attaches them to the objective function

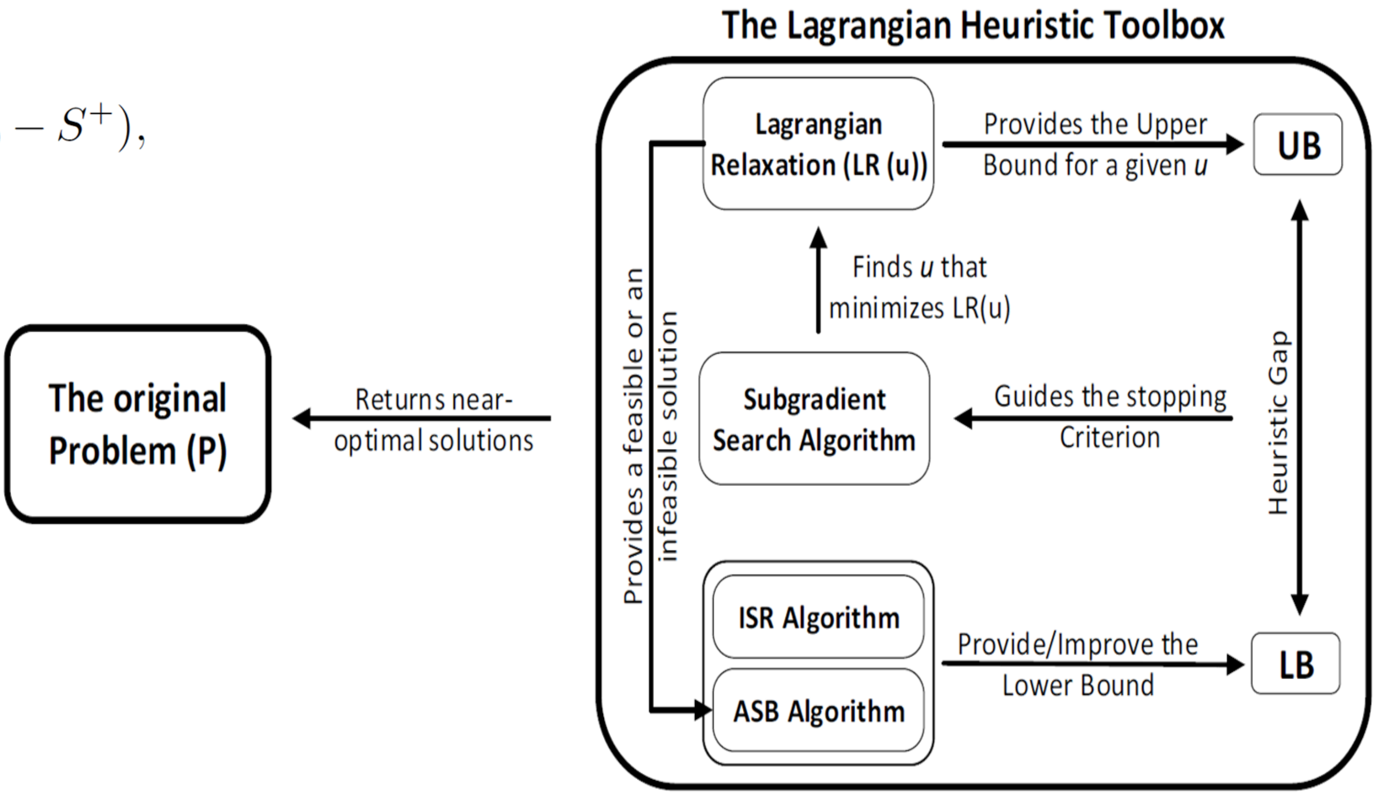
$$(LR_u) \qquad \max Z^{LR_u} = \sum_{i=1}^{N} \sum_{t=0}^{T} [(X_{it} - Y_{it}) + u(\sum_{i=1}^{N} X_{i0} - S^{+}),$$

- Solves the Lagrangian Dual problem to find the optimal Lagrangian multipliers

$$(LD_u) \qquad Z^{LD_u}(u) = Min_u Z^{*LR_u}(u),$$

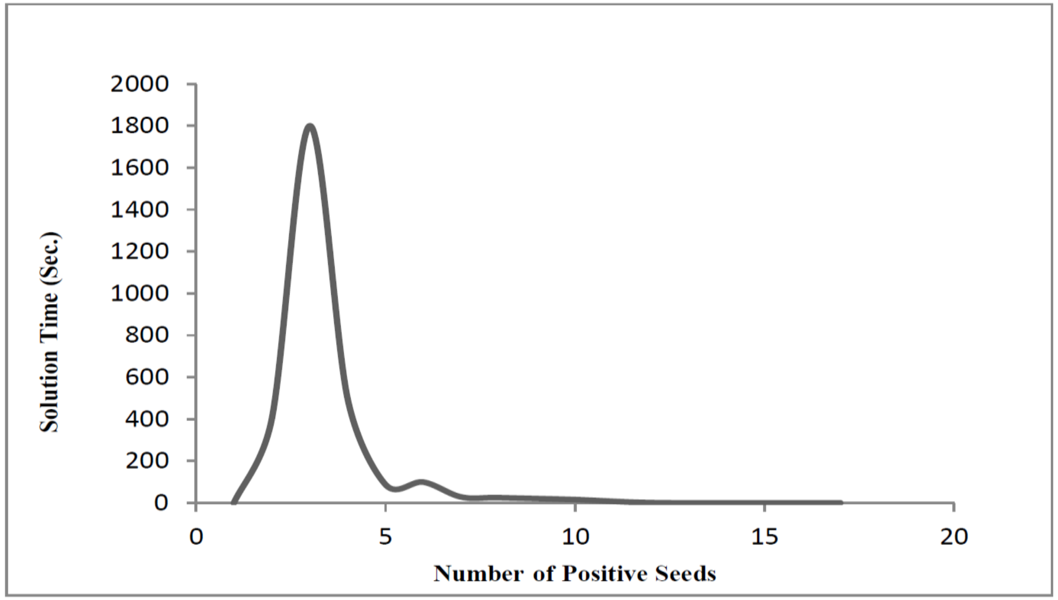
- Uses Subgradient search algorithm for solving the Lagrangian Dual problem.

The introduced Influence Maximization problem is NP-hard.


The Lagrangian Heuristic Toolbox
Lagrangian Relaxation (LR (u))
Provides the Upper Bound for a given u
UB
Finds u that minimizes LR(u)
Subgradient Search Algorithm
Guides the stopping Criterion
Heuristic Gap
The original Problem (P)
Returns near-optimal solutions
Provides a feasible or an infeasible solution
ISR Algorithm
ASB Algorithm
Provide/Improve the Lower Bound
LB

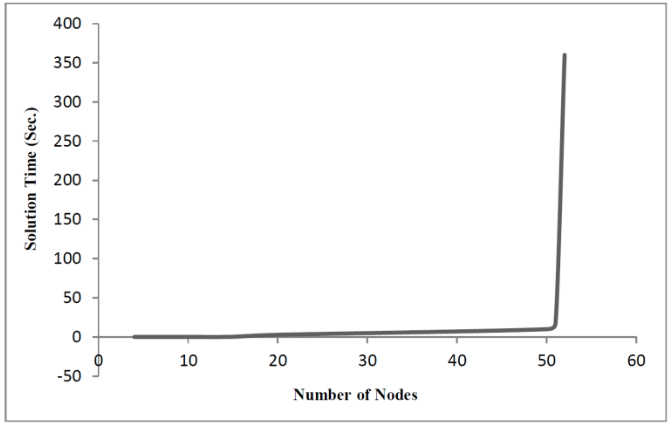
The Lagrangian heuristic toolbox: an overview of the components.

Sources of Complexity



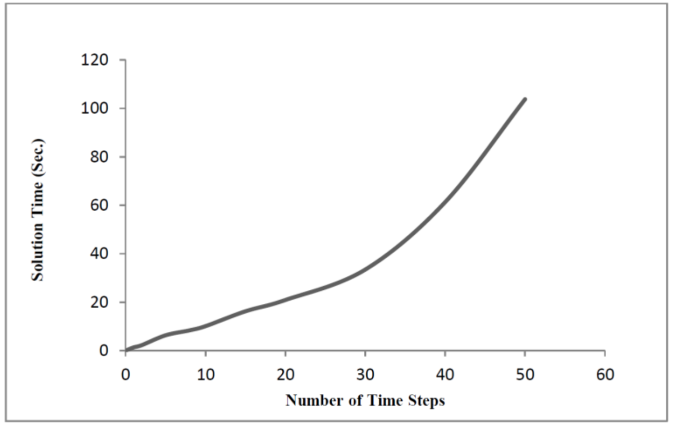
Observation 1.

- Solution time first increases with the number of positive seeds and then decreases.
- This diagram shows why the solution time for dummy problem is significantly lower than (P).
- Serves as the basic idea of ISR algorithm



Observation 2.

- Solution time exponentially increases with number of nodes



Observation 3.

- Solution time smoothly increases with the number of time periods.

Finding Lower Bound

- Two heuristics are developed for finding lower bound for the optimal solution and stopping the search procedure.

Iterative Seed Removal (ISR) Algorithm

Finding a dummy problem with more positive seeds

The solution time for dummy problem is significantly lower than (P).

Solution of the dummy problem is expected to include the original problem's Solution.

ISR iteratively removes the seeds in dummy solution to provide a valid lower bound for (P).

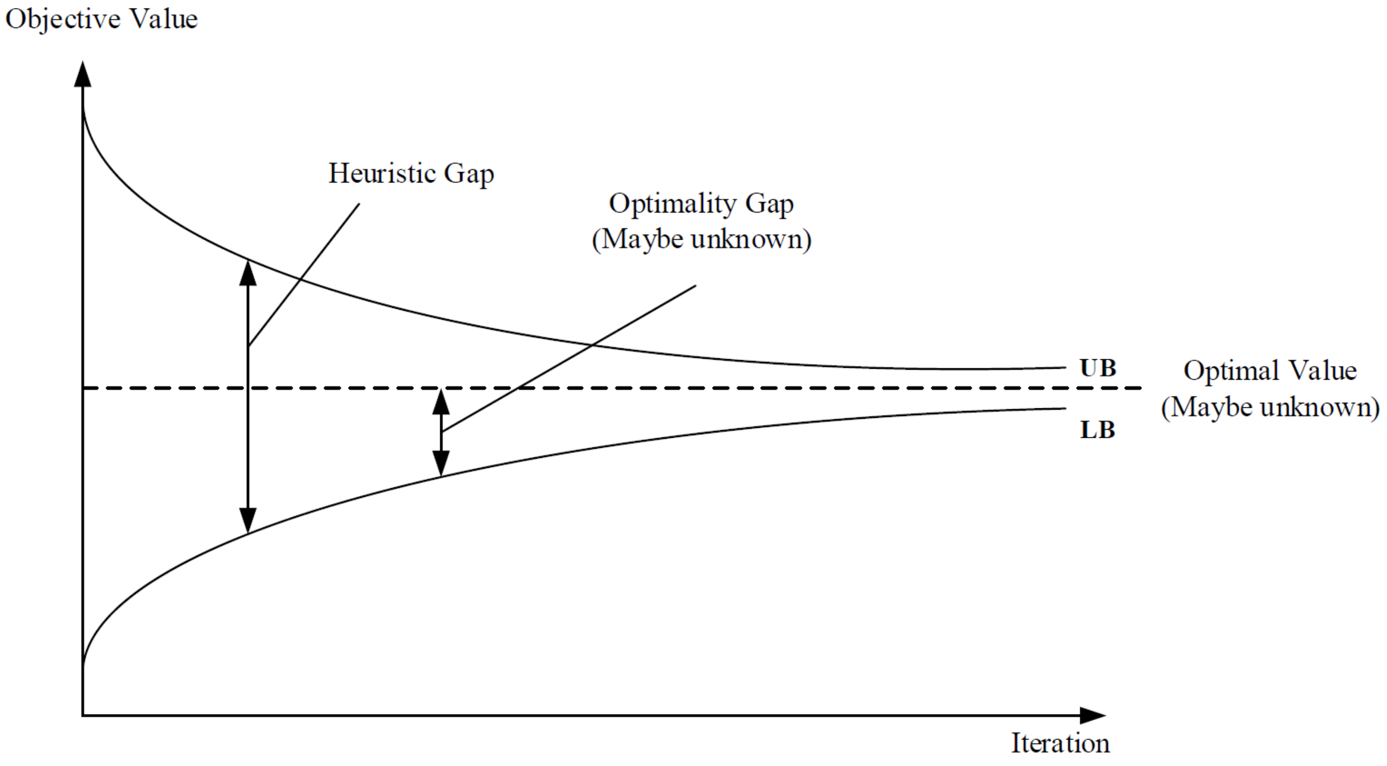
Adaptive Subgradient-Based (ASB) Algorithm

ASB utilizes the information in subgradient algorithm.

In each iteration, the Lagrangian Relaxation problem returns a solution with more positive seeds.

ASB selects the first k_1 positive seeds from the solution of the relaxed problem.

Heuristic and Optimality Gap



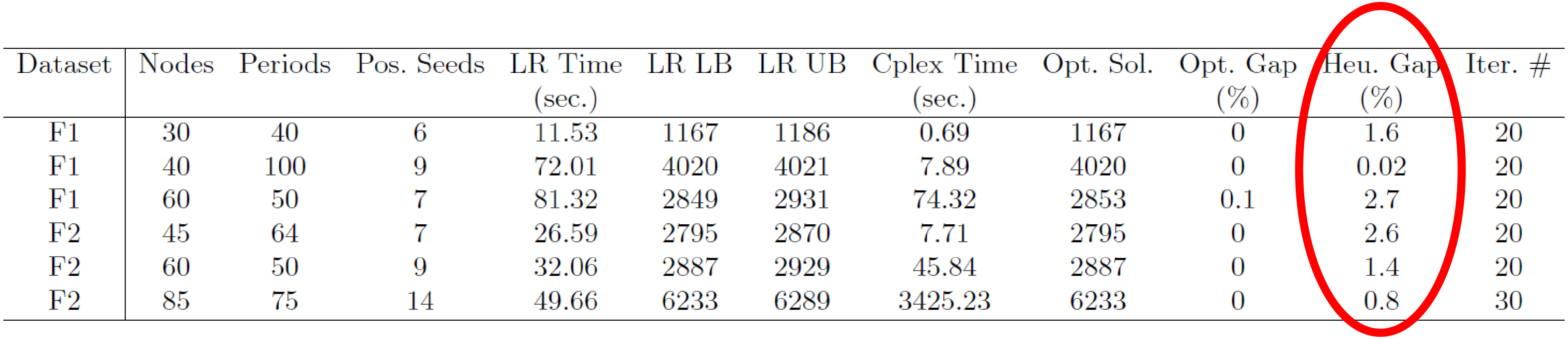
- The lower bound is obtained by two heuristic methods.
- Upper bound is obtained by Lagrangian Relaxation.
- The Lagrangian Relaxation heuristic guarantees the quality of the solution.

Computational Results (Small- sized problems)

- For the small problems, CPLEX outperforms the Lagrangian Relaxation heuristic in terms of solution time.

- When the problem size increases, the solution time of CPLEX increases rapidly but the Lagrangian Relaxation heuristic remains fast.

Max Gap =2.7%

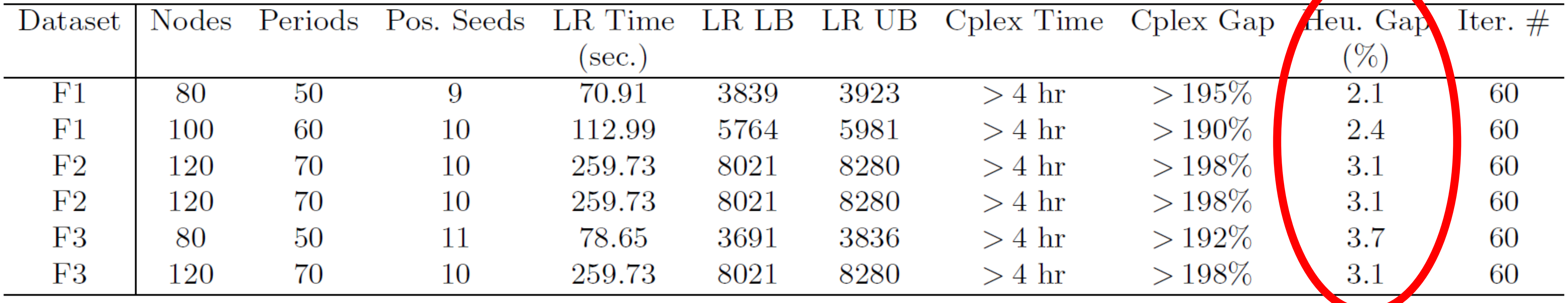


Dataset	Nodes	Periods	Pos. Seeds	LR Time (sec.)	LR LB	LR UB	Cplex Time (sec.)	Opt. Sol.	Opt. Gap (%)	Heu. Gap (%)	Iter. #
F1	30	40	6	11.53	1167	1186	0.69	1167	0	1.6	20
F1	40	100	9	72.01	4020	4021	7.89	4020	0	0.02	20
F1	60	50	7	81.32	2849	2931	74.32	2853	0.1	2.7	20
F2	45	64	7	26.59	2795	2870	7.71	2795	0	2.6	20
F2	60	50	9	32.06	2887	2929	45.84	2887	0	1.4	20
F2	85	75	14	49.66	6233	6289	3425.23	6233	0	0.8	30

Computational Results (Mid-Sized Problems)

- For mid-sized Facebook networks, CPLEX cannot even create a feasible solution in the computer memory.
- The Lagrangian Relaxation heuristic runs in a reasonable computational time and provides an acceptable heuristic gap.
- The runtime for the Lagrangian Relaxation heuristic smoothly increases with the dimensions of the problem instances

Max Gap =3.7%



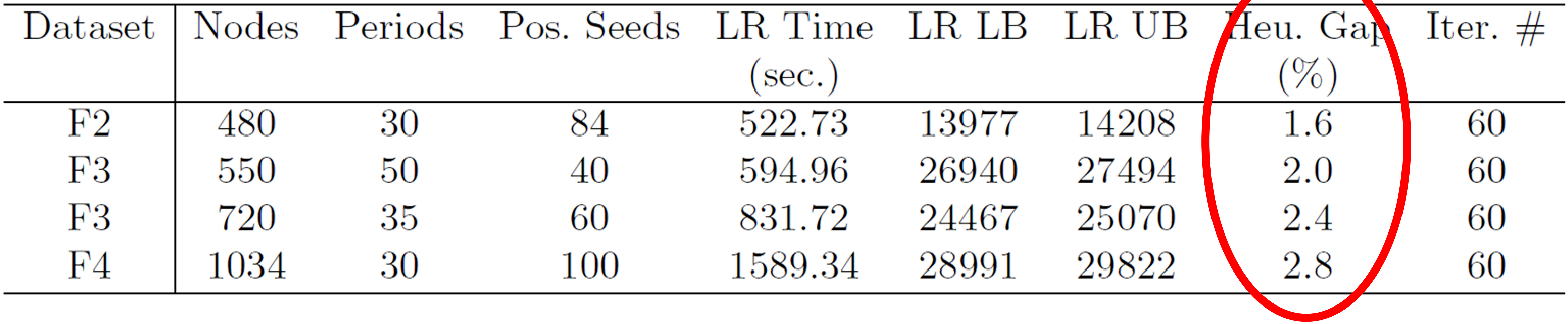
Dataset	Nodes	Periods	Pos. Seeds	LR Time (sec.)	LR LB	LR UB	Cplex Time	Cplex Gap	Heu. Gap (%)	Iter. #
F1	80	50	9	70.91	3839	3923	> 4 hr	> 195%	2.1	60
F1	100	60	10	112.99	5764	5981	> 4 hr	> 190%	2.4	60
F2	120	70	10	259.73	8021	8280	> 4 hr	> 198%	3.1	60
F2	120	70	10	259.73	8021	8280	> 4 hr	> 198%	3.1	60
F3	80	50	11	78.65	3691	3836	> 4 hr	> 192%	3.7	60
F3	120	70	10	259.73	8021	8280	> 4 hr	> 198%	3.1	60

Computational Results (Large-Sized Problems)

- For large-size problems, we don’t try CPLEX.

- The solution time and the heuristic gap are still acceptable.

Max Gap =2.8%



Dataset	Nodes	Periods	Pos. Seeds	LR Time (sec.)	LR LB	LR UB	Heu. Gap (%)	Iter. #
F2	480	30	84	522.73	13977	14208	1.6	60
F3	550	50	40	594.96	26940	27494	2.0	60
F3	720	35	60	831.72	24467	25070	2.4	60
F4	1034	30	100	1589.34	28991	29822	2.8	60

Seed Selection Scheduling

- Relaxing the constraint on seed activation time.
- Increasing the solution space from $\binom{k}{N}$ to $\binom{kT}{NT}$.
- Initiating the cascade with a portion of seeds and saving others as reminders.
- The idea works when forgetfulness parameter β and evidence transfer efficiency reduction parameter α exist in the model.
- The results confirm that with $\beta < 1$ and $\alpha < 1$, the whole network become positively activated and then lose their activation.
- Late seed activation delays the deactivation time.

Seed Selection Scheduling

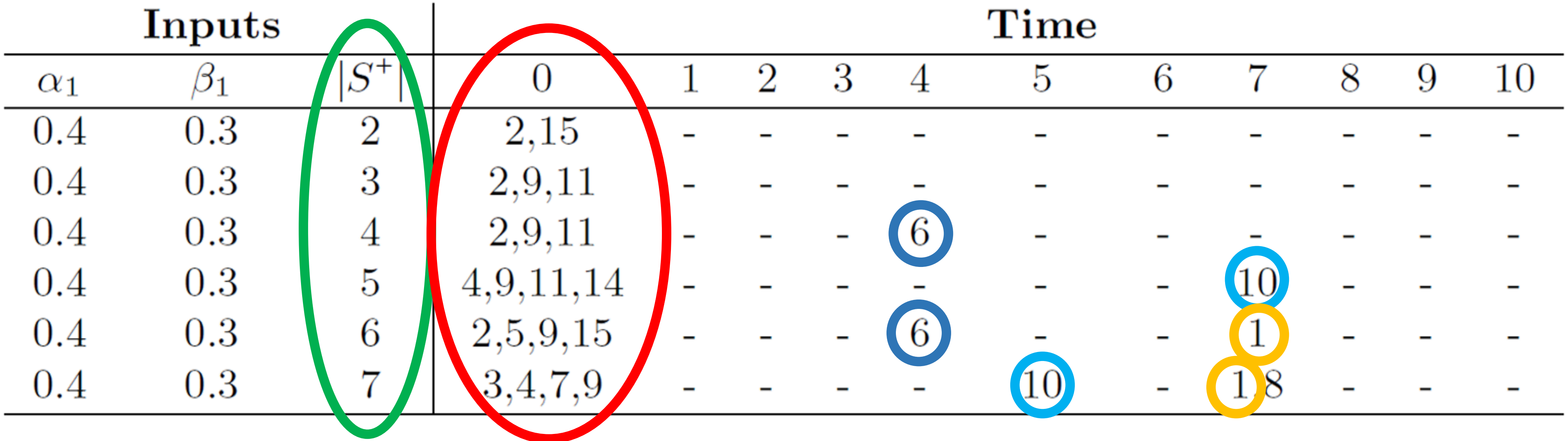
- The designed PCIM for the regular influence maximization does not work for seed selection scheduling.
- The regular modeling does not allow selecting negatively activated nodes as positive seeds.
- The seed activation is changed so that the positive seeds receive a package of positive evidence.
- This activation includes the situation that selected positive seed does not necessarily get the positive activation.

$$L_{it} = \beta_1 L_{it-1} + \sum_{(j,i)\in A} E^{+}_{jt-1} + S^{+}_{it} \boxed{q^{+}} \qquad i = 1, 2, \ldots |N|, \quad t = 1, 2, \ldots, T,$$

Positive evidence package

Seed Selection Scheduling (Results)

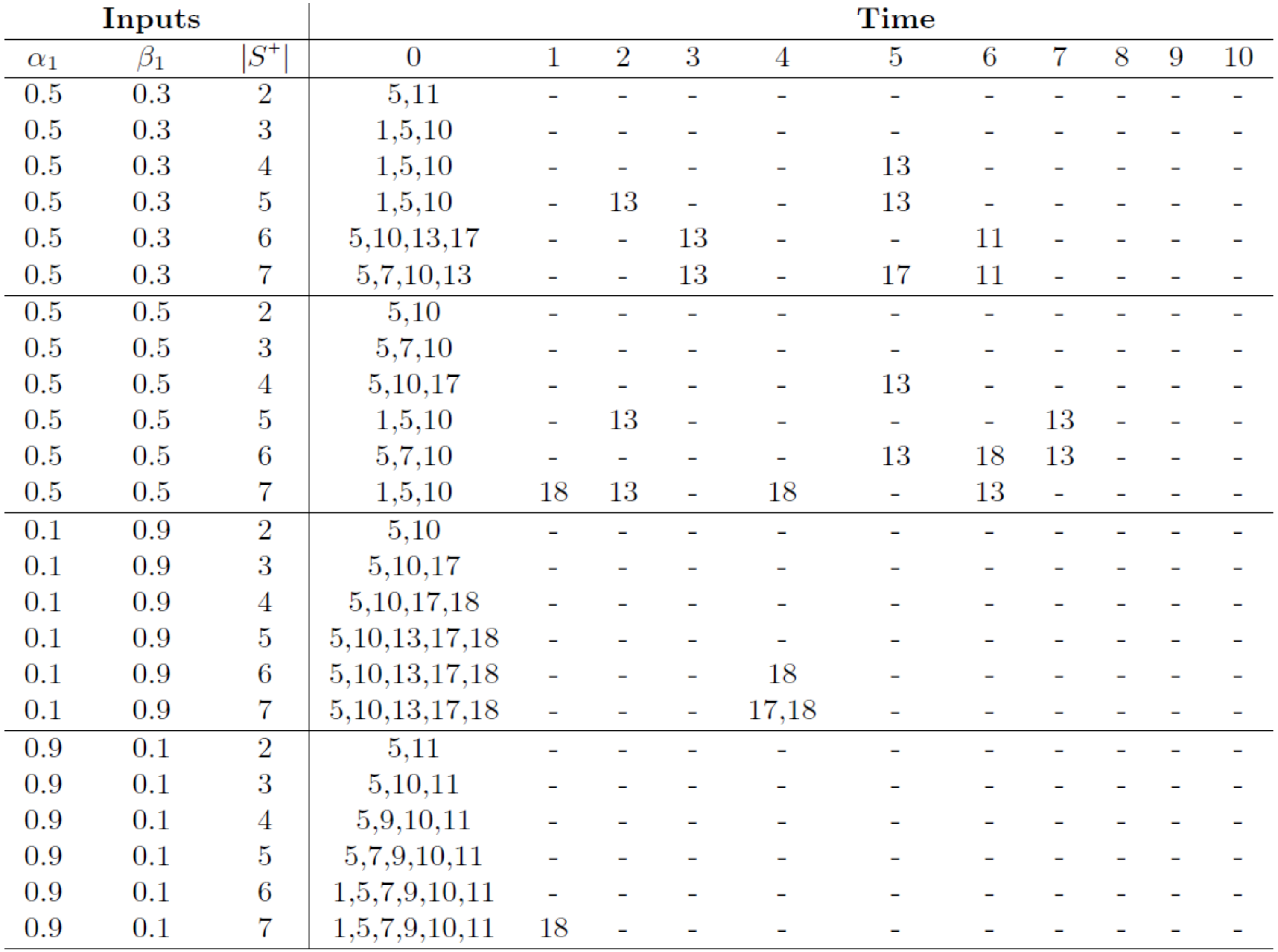
- As the number of positive seeds (budget) increases, the decision-maker considers late seed activation on Florentine families' marriage network.



Inputs			Time										
α_1	β_1	$\|S^+\|$	0	1	2	3	4	5	6	7	8	9	10
0.4	0.3	2	2,15	-	-	-	-	-	-	-	-	-	-
0.4	0.3	3	2,9,11	-	-	-	-	-	-	-	-	-	-
0.4	0.3	4	2,9,11	-	-	-	6	-	-	-	-	-	-
0.4	0.3	5	4,9,11,14	-	-	-	-	-	-	10	-	-	-
0.4	0.3	6	2,5,9,15	-	-	-	6	-	-	1	-	-	-
0.4	0.3	7	3,4,7,9	-	-	-	-	10	-	1,8	-	-	-

- Fixed pattern of late seed activation! All late seeds close to negative seed.

Seed Selection Scheduling (Results)

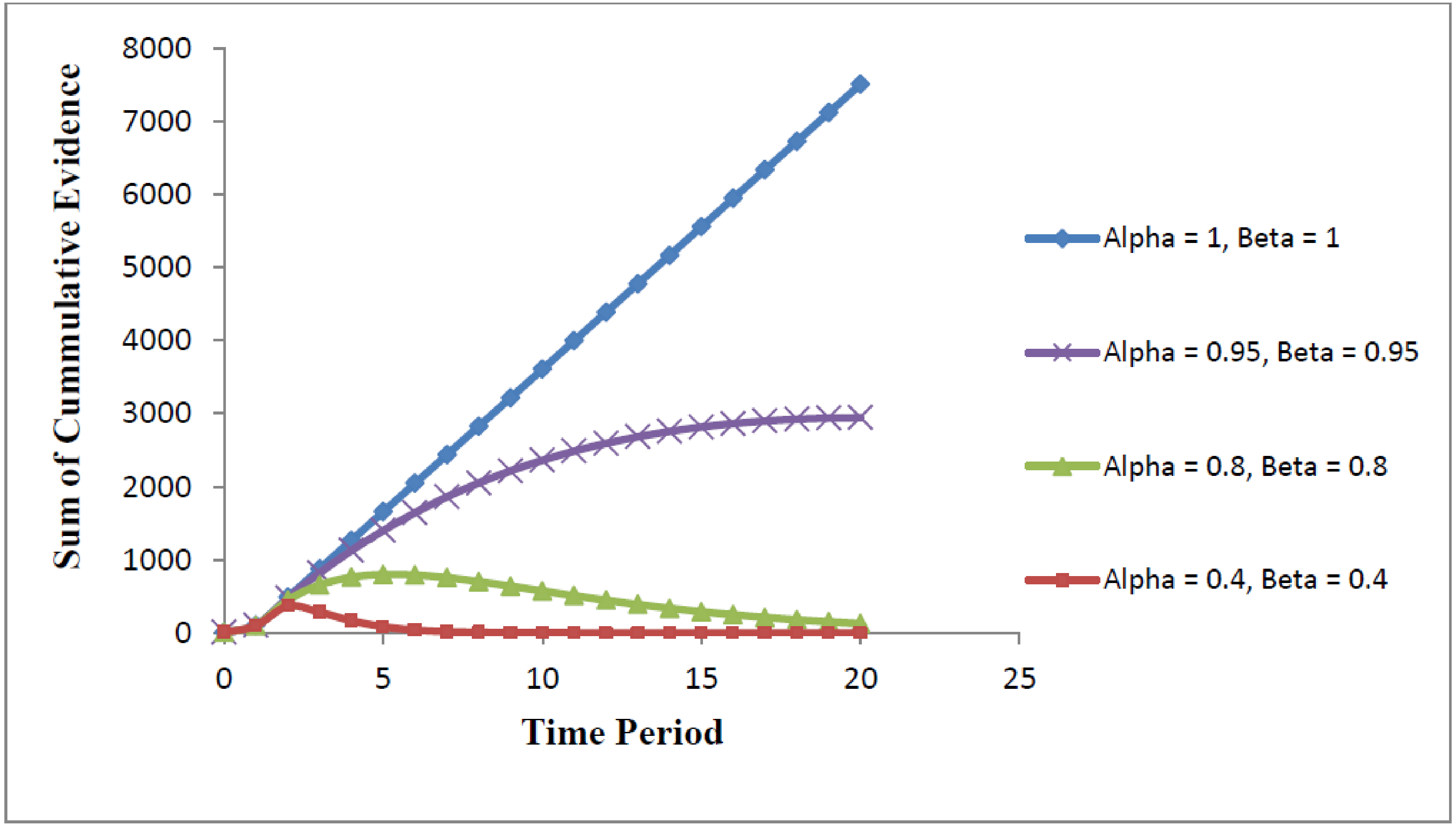


Inputs			Time										
α_1	β_1	$\lvert S^+ \rvert$	0	1	2	3	4	5	6	7	8	9	10
0.5	0.3	2	5,11	-	-	-	-	-	-	-	-	-	-
0.5	0.3	3	1,5,10	-	-	-	-	-	-	-	-	-	-
0.5	0.3	4	1,5,10	-	-	-	-	13	-	-	-	-	-
0.5	0.3	5	1,5,10	-	13	-	-	13	-	-	-	-	-
0.5	0.3	6	5,10,13,17	-	-	13	-	-	11	-	-	-	-
0.5	0.3	7	5,7,10,13	-	-	13	-	17	11	-	-	-	-
0.5	0.5	2	5,10	-	-	-	-	-	-	-	-	-	-
0.5	0.5	3	5,7,10	-	-	-	-	-	-	-	-	-	-
0.5	0.5	4	5,10,17	-	-	-	-	13	-	-	-	-	-
0.5	0.5	5	1,5,10	-	13	-	-	-	-	13	-	-	-
0.5	0.5	6	5,7,10	-	-	-	-	13	18	13	-	-	-
0.5	0.5	7	1,5,10	18	13	-	18	-	13	-	-	-	-
0.1	0.9	2	5,10	-	-	-	-	-	-	-	-	-	-
0.1	0.9	3	5,10,17	-	-	-	-	-	-	-	-	-	-
0.1	0.9	4	5,10,17,18	-	-	-	-	-	-	-	-	-	-
0.1	0.9	5	5,10,13,17,18	-	-	-	-	-	-	-	-	-	-
0.1	0.9	6	5,10,13,17,18	-	-	-	18	-	-	-	-	-	-
0.1	0.9	7	5,10,13,17,18	-	-	-	17,18	-	-	-	-	-	-
0.9	0.1	2	5,11	-	-	-	-	-	-	-	-	-	-
0.9	0.1	3	5,10,11	-	-	-	-	-	-	-	-	-	-
0.9	0.1	4	5,9,10,11	-	-	-	-	-	-	-	-	-	-
0.9	0.1	5	5,7,9,10,11	-	-	-	-	-	-	-	-	-	-
0.9	0.1	6	1,5,7,9,10,11	-	-	-	-	-	-	-	-	-	-
0.9	0.1	7	1,5,7,9,10,11	18	-	-	-	-	-	-	-	-	-

- On a Twitter network, increasing α makes the decision-maker to select initial seeds.

- With a large α, the cascade has a chance to remain alive for a long time if e^+ is large enough.

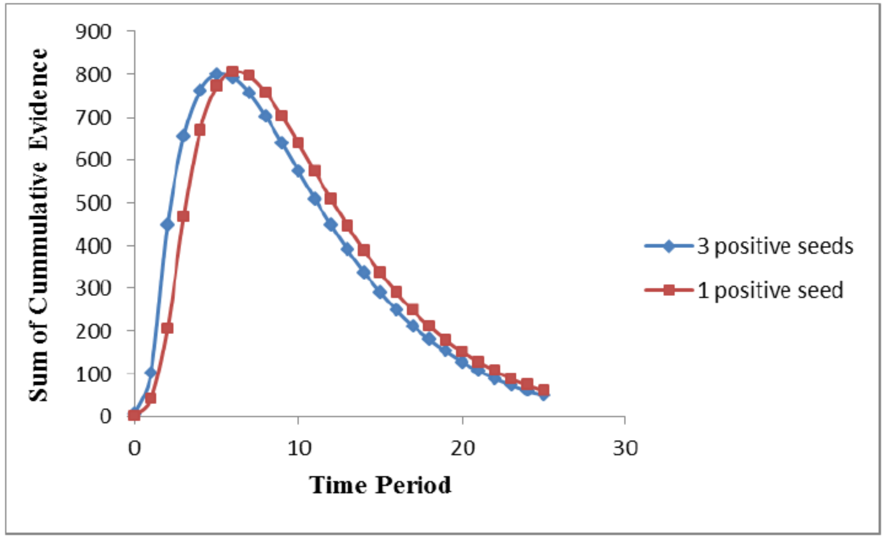
Cumulative evidence spread through the network

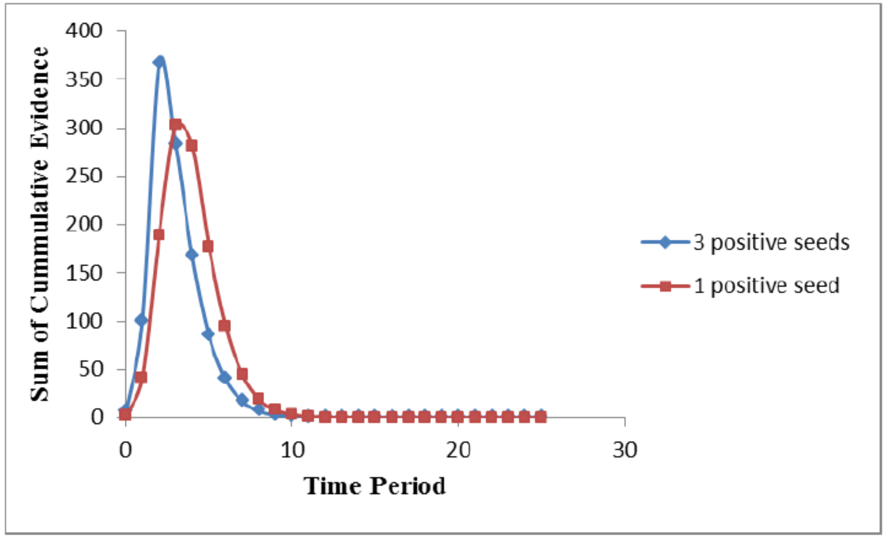


- The sum of cumulative evidence on Zachary's karate club network is concave if $\beta < 1$ and $\alpha < 1$.

- Decreasing α and β delays the convergence (long-term steady state).

Cumulative evidence spread through the network



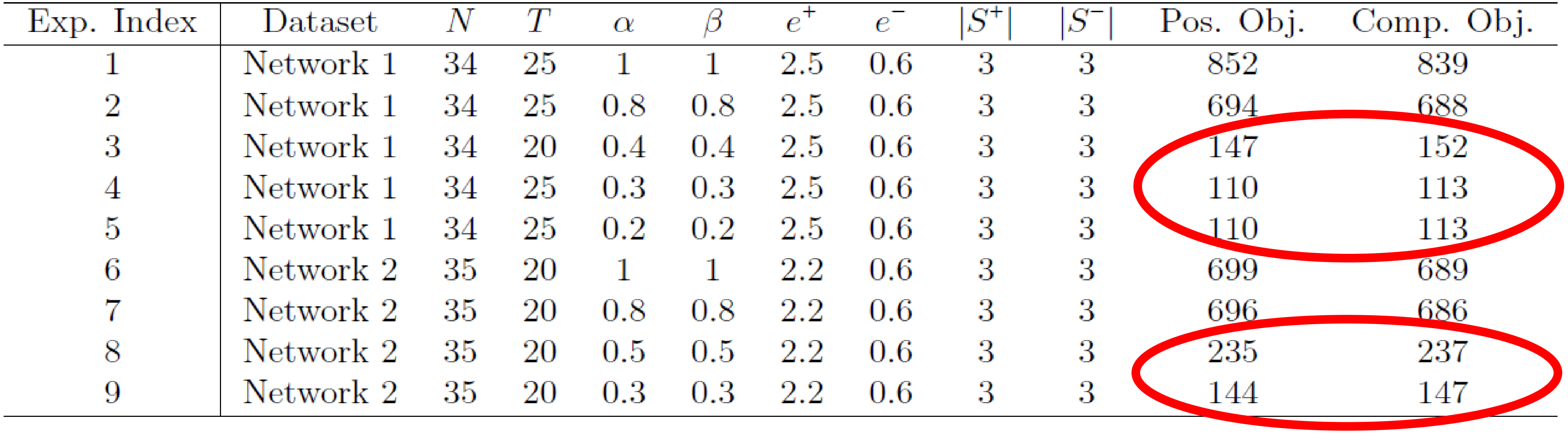


- In a pure positive cascade, increasing the number of positive seeds delays the long-term steady state.

- When the decision-maker has enough budget to initiate a strong cascade, it might be better to start the cascade slowly!

Positive evidence persistency

- **Observation:** The presence of the negative evidence makes the positive evidence more consistent (helps it stay alive longer) in the network and delays the long-term steady state.

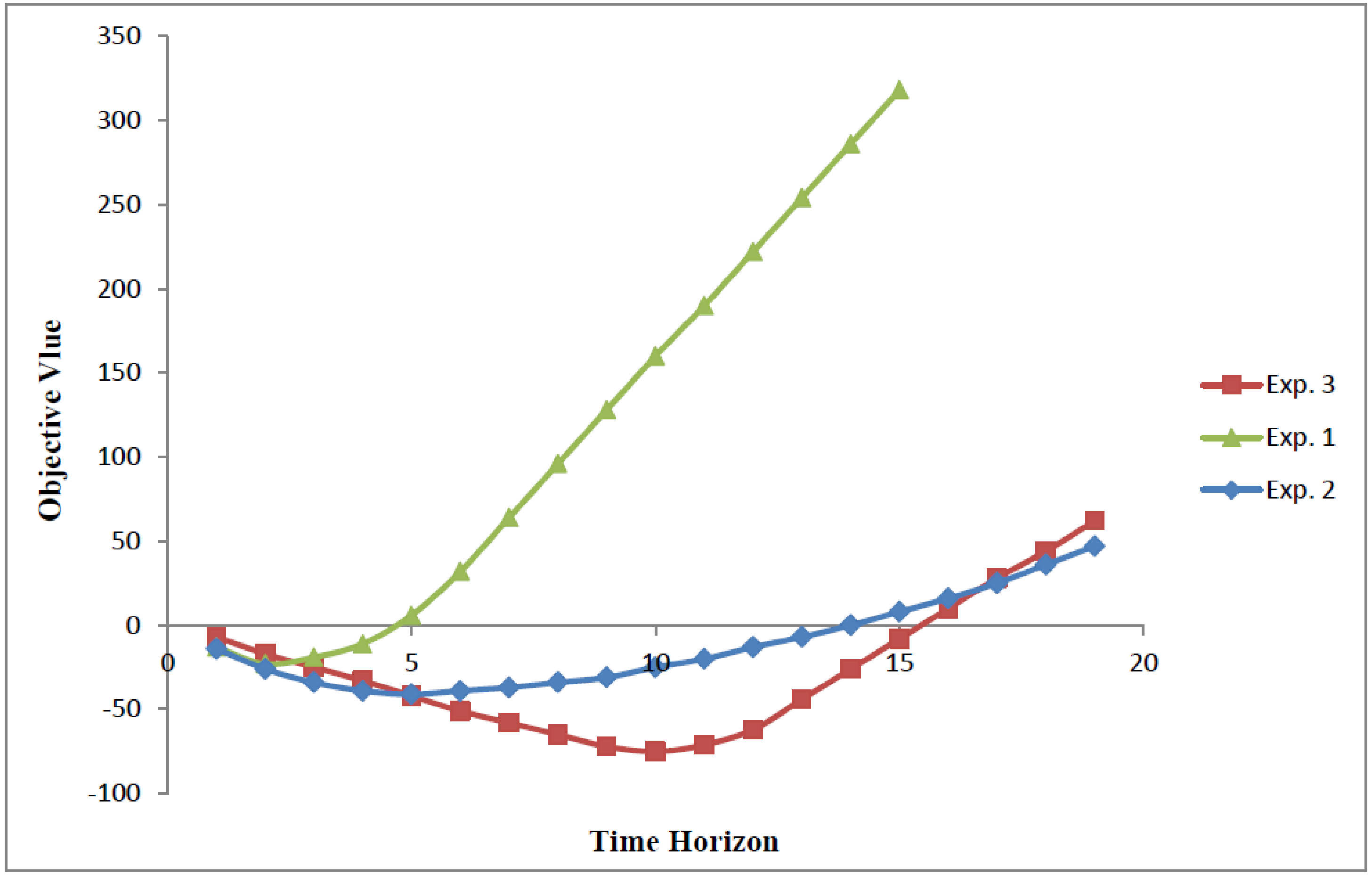


Exp. Index	Dataset	N	T	α	β	e^+	e^-	$\|S^+\|$	$\|S^-\|$	Pos. Obj.	Comp. Obj.
1	Network 1	34	25	1	1	2.5	0.6	3	3	852	839
2	Network 1	34	25	0.8	0.8	2.5	0.6	3	3	694	688
3	Network 1	34	20	0.4	0.4	2.5	0.6	3	3	147	152
4	Network 1	34	25	0.3	0.3	2.5	0.6	3	3	110	113
5	Network 1	34	25	0.2	0.2	2.5	0.6	3	3	110	113
6	Network 2	35	20	1	1	2.2	0.6	3	3	699	689
7	Network 2	35	20	0.8	0.8	2.2	0.6	3	3	696	686
8	Network 2	35	20	0.5	0.5	2.2	0.6	3	3	235	237
9	Network 2	35	20	0.3	0.3	2.2	0.6	3	3	144	147

Managerial insight

- For a strongly dominant in a social network, it is better to let the weak competitors remain in the market.
- The competition makes the positive evidence more persistent.
- A pure positive cascade without competition eventually disappears.

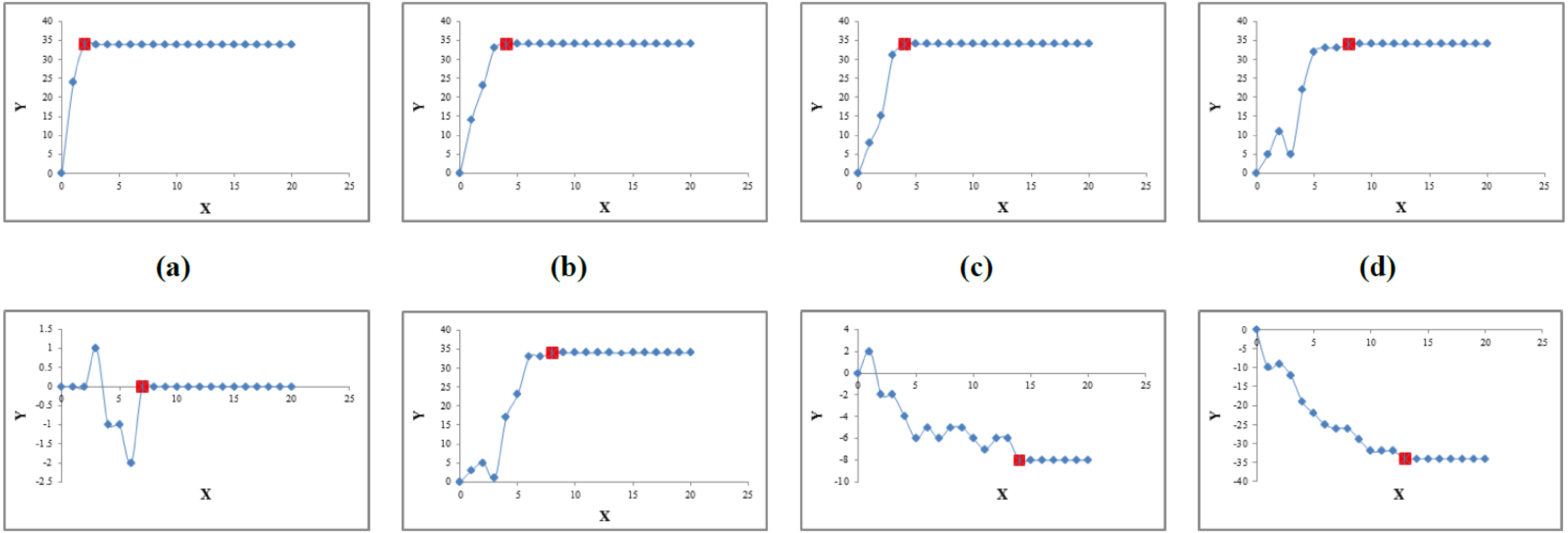
Close competition on network



- **Observation 2:** In a close competition with negative gain, the decision-maker can make the cascade positive if the time horizon is large enough.
- **Game:** The negative takes the action first. The positive takes the action while knows the negative action.
- In a close competition, long time horizon is better for the positive party!

- Time horizon to get stable seed selection strategies depends on evidence increments.

X: Time horizon **Y:** Marginal change of the objective value



(a) (b) (c) (d)

Time till stability

- Running for a very large time horizon is computationally expensive and gets the same seed selection strategy in the stability.

- Running for a very small time horizon (before stability) reduces the computations but does not guarantee a long-term optimal seed selection strategy.

- It takes longer for a negatively dominant cascade to reach the stability.

- The positive party is somehow the leader in the game.

Future Research

- Future studies can apply the proposed optimization scheme for modeling the spread of evidence in growing social networks.
- Further research is required to employ network-level metrics, e.g., clustering coefficient, for reducing the size of large influence maximization problems to make them manageable.

Thank you

This research has been supported in part by the National Science Foundation (Award #62288) and by a Multidisciplinary University Research Initiative (MURI) grant (#W911NF-09-1-0392) for “Unified Research on Network-based Hard/Soft Information Fusion”, issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery.