Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

SILCOM: an open framework for reconstructing, simulating and mitigating withdrawal cascades on Bluesky

Mousa Al-Kfairy

Mousa.Al-kfairy@zu.ac.ae

Zayed University

Research Article

Keywords: Bluesky, Decentralized social networks, Withdrawal cascades, Social contagion, Agent-based modeling, Churn prediction

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11661100/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

SILCOM: an open framework for reconstructing, simulating and mitigating withdrawal cascades on Bluesky

Mousa Al-Kfairy[1*]

[1*]College of Technological Innovation, Zayed University, Abu Dhabi, United Arab Emirates.

Corresponding author(s). E-mail(s): Mousa.Al-kfairy@zu.ac.ae;

Abstract

Online communities last only while their members keep taking part. When central or supportive members leave, others may follow. Many communities have moved to decentralized platforms such as Bluesky. Yet no open tool lets researchers rebuild how such a community has been losing members, test why, and try out responses first. We present SILCOM, an open framework built on one retrospective collection of public data from Bluesky's AT Protocol. It rebuilds participation histories and tests whether departures cluster in the interaction network. It fits seven competing withdrawal mechanisms in an agent-based model and selects among them on held-out months. It predicts who is at risk and compares interventions at a fixed budget. A synthetic benchmark with planted mechanisms supplies ground truth. We apply SILCOM to 3,121 Bluesky accounts over 18 months and compare every module with published alternatives. Departures cluster beyond chance (permutation $\boldsymbol{p = 0.005}$; shuffle test $\boldsymbol{p = 0.04}$). The benchmark shows, however, that no test, ours included, reliably separates contagion from homophily. The independent-cascade model fitted held-out months best. Under it, removing the 5% best-connected users caused extra departures equal to 49% of the shock. SILCOM's risk model matched a graph-convolution baseline (AUC 0.83) and beat survival, $\boldsymbol{k}$-core and exposure rules. Support aimed at predicted risk kept 3.5 times more users than random support. Aiming it by degree, betweenness, $\boldsymbol{k}$-shell, PageRank or collective influence kept almost none. We close with implications for community organizers and platform designers and an agenda for the social network analysis and mining community.

Keywords: Bluesky, Decentralized social networks, Withdrawal cascades, Social contagion, Agent-based modeling, Churn prediction

1 Introduction

Online communities need a steady flow of contributions, and most struggle to keep it. In six public newsgroups, most people who posted once never posted again (Joyce and Kraut, 2006). Newcomers who got a reply were 12% more likely to return. Communities last longer when communication is responsive and deep (Kang and Shin, 2006; Petrič, 2015). A sense of belonging helps (Bessems et al, 2022), as does attachment to subgroups (Tausczik et al, 2014). A community rests on the ties between its members, and the platform only hosts them (Stanoevska-Slabeva and Schmid, 2001).

Because members are linked, they may not leave independently. In a mobile phone network, people were more likely to switch provider when more of their contacts had switched (Dasgupta et al, 2008). In an online game, churn clustered in the network and reached three steps away (Ding et al, 2015). In online social networks, whether a user leaves depends on how many friends stay active (Wu et al, 2013). Friendster's collapse has been read as a cascade (Garcia et al, 2013). When staying cost more than it gave, users left. Their friends then got less from staying, and some left too. Structure, then, may shape how much of a community survives a shock.

This now matters for decentralized platforms. After Twitter changed owners in 2022, many users moved to Bluesky, Mastodon and similar services (Jeong et al, 2024; Radivojevic et al, 2025; Ng and Ray, 2025). The move spread through social ties (La Cava et al, 2023), and scholars describe it as a collective event (Tkacz et al, 2026; Ittefaq, 2025). Bluesky splits a platform into parts that different providers can run, on the open AT Protocol (Kleppmann et al, 2024; Balduf et al, 2024). Every account keeps a public, signed repository of its records. Whether communities formed there will last is unclear (Struett et al, 2024). Groups that arrived together may also leave together.

The tools to study this are missing. Collapse models have been fitted after the fact to single, centralized platforms (Garcia et al, 2013). Robustness studies remove nodes but ignore how the rest react (Tanizawa et al, 2005; Trajanovski et al, 2013). Measurement studies of Bluesky describe structure, activity and moderation (Quelle and Bovet, 2025; Nogara et al, 2026; Sokoto et al, 2026). They do not model decline. Churn models that use social ties run on private data from one service (Oentaryo et al, 2012; Zhang et al, 2022; Lu et al, 2026). They predict churn but cannot test responses to it. Agent-based models of online behavior are seldom fitted to data (Boero and Squazzoni, 2005; Sittar et al, 2026). Observational data also struggle to separate contagion from homophily and shared shocks (Shalizi and Thomas, 2011; O'Malley et al, 2014).

Timing makes the gap pressing. Several large centralized platforms have restricted research access to their data. Bluesky does the opposite: every account's history can be read from public repositories, after the fact and without special access. For the first time, a researcher can rebuild a whole community's participation history from public data and test how it declines. The social network analysis and mining community, with its tools for contagion, robustness and churn, is well placed to use this opening.

We present Silcom (*Silent Community*), an open framework for this task. It needs one retrospective collection of public data, so it can be reused on any community without months of monitoring. It rebuilds monthly histories and tests whether departures

cluster in the network. It fits competing withdrawal mechanisms in an agent-based model, including two classic diffusion models (Kempe et al, 2003), and selects among them on held-out months. It predicts who is at risk and compares interventions on a fixed budget. A synthetic benchmark with planted mechanisms provides ground truth. Five research questions guide the evaluation:

RQ1 (Reconstruction) How fully and accurately can one retrospective collection rebuild participation, interaction and withdrawal on Bluesky?
RQ2 (Mechanisms) Can the framework tell withdrawal mechanisms apart when the truth is known? Which mechanism best predicts held-out real data?
RQ3 (Risk) How well does it identify users at risk?
RQ4 (Comparison) How do its modules compare with published methods for detecting contagion, predicting churn, modeling cascades and targeting interventions?
RQ5 (Utility) Is losing hubs or supportive users worse than random loss? Which intervention, and which way of targeting it, keeps the most users for a fixed budget?

1.1 Contributions

This article offers four contributions.

1. **An open system for decentralized platforms.** SILCOM is an open pipeline from public Bluesky data to a calibrated simulation and an intervention testbed. It reads AT Protocol repositories directly, rebuilds histories after the fact and scales near-linearly to 50,000 users (Sections 3 and 4).
2. **Methods with ground truth.** A calibration, validation and selection procedure over seven withdrawal mechanisms; a paired design that measures cascades with common random numbers; a risk prediction module; and a synthetic benchmark with planted mechanisms that shows which conclusions the framework can support (Sections 3–4).
3. **Head-to-head comparison with published methods.** Every module is compared with established alternatives for contagion testing, cascade modeling, churn prediction and intervention targeting, on the same data and splits. The comparison shows where the framework is reliable and where no current method is (Section 5).
4. **Findings, implications and agenda.** Departures on Bluesky cluster in the network, and a standard cascade model fits them best. Losing hubs is amplified most. Whom an intervention targets matters more than what it does, and network centrality is the wrong target. We draw out implications for community organizers and platform designers and set an agenda for decentralized networks (Sections 6–10).

Prior work stays within one stream. It explains community decline, detects social contagion, predicts churn or measures decentralized platforms. Our contribution joins these streams in one validated system and tests each part against the methods of its own stream. Table 1 makes the difference explicit.

The paper proceeds as follows. Section 2 reviews related work. Section 3 describes SILCOM. Section 4 evaluates it (RQ1–RQ3). Section 5 compares it with published methods (RQ4). Section 6 reports shocks and interventions (RQ5). Section 7 discusses the findings, Section 8 situates them in computational social science, Section 9 draws

out practical implications and Section 10 sets a research agenda. Sections 11–12 cover limitations and conclusions.

2 Related work

We draw on six bodies of work. They cover why online communities decline, how departures spread, how networks break and whom to target, how churn is predicted, what is known about decentralized platforms, and how agent-based models are tied to data. Table 1 sums up what each offers and what it leaves open.

2.1 Sustainability and decline of online communities

Why do a few online communities thrive while most fade? Early newsgroup studies found that a reply to a first post predicts a return (Joyce and Kraut, 2006). Responsive, broad and deep communication goes with member loyalty (Kang and Shin, 2006). Surveys point to the quality of communication (Petrič, 2015), social support and identity (Chiu et al, 2015), and a sense of belonging (Bessems et al, 2022). Features that signal trust also raise participation (Benlian et al, 2010; Benlian and Hess, 2011). In experiments, talking within sub-groups made people more loyal to the whole community (Tausczik et al, 2014). Offline meetups raised Wikipedia contributions in the short and long run (Schwitter, 2024).

Network structure matters at the community level too. Loyal Reddit communities have denser interaction networks and less triadic closure (Hamilton et al, 2017). They hold together instead of splitting into cliques. Density, clustering, reciprocity and centralization affect whether Reddit communities survive, though barely whether they grow (Tsugawa and Niida, 2019). Adding such network features raised the F1 score of survival prediction by about 30% (Tsugawa and Niida, 2021). Other work predicts whether communities will grow, shrink or dissolve from structural and temporal features (Diakidis et al, 2015; Pavlopoulou et al, 2017). Within a community, a few highly engaged users often carry much of the talk (Myneni et al, 2018). Losing them may hurt more than losing others.

Web research has followed communities and sites through their whole life cycle. Whether people join a group depends on how their friends inside it are connected (Backstrom et al, 2006). Group growth and longevity can be predicted from diffusion and structural features (Kairam et al, 2012). Users who stop adapting to a community's changing language tend to leave soon after (Danescu-Niculescu-Mizil et al, 2013). Early activity predicts whether new users of a question-answering site churn (Dror et al, 2012). Whole membership-based websites grow or die depending on how active members recruit and keep others (Ribeiro, 2014). In Stack Exchange communities, activity rises and decays as a dynamical process driven by user interaction (Walk et al, 2016).

These studies show *which* communities decline. They rarely model *how* decline unfolds. The cost–benefit model of Garcia et al (2013) is an exception. It explains the decline of Friendster, Livejournal, Facebook, Orkut and Myspace through each network's k-core distribution. Its authors conclude that topology alone cannot explain

success or failure. The model was fitted after the fact to centralized platforms, and it assumes one fixed rule for leaving.

2.2 Social contagion of disengagement

Departures tend to cluster among connected users. In a mobile network, churn spread along social ties, so current churners helped predict future ones (Dasgupta et al, 2008). In an online game, active and inactive players formed visible clusters (Ding et al, 2015). Friends with more common friends were more alike in leaving. In a large online network, users with few friends left according to how many stayed active (Wu et al, 2013). Users with many friends left according to the share that stayed. Departures were correlated in tight-knit groups. A structural model of another game found that peer effects amplify retention efforts (Chen and Nevskaya, 2026). Users differed more in how they affect others' retention than in their own chance of staying.

Social support looks like a key channel. In an online health community, social capital and support reduced churn (Zhang and Lu, 2026). Engagement alone did not, unless it came with support. In another health community, likes and comments received were both linked to longer retention (Wan et al, 2020). Generalized reciprocity keeps peer-support communities going (Phan and Yarosh, 2016). Losing reciprocal partners, more than losing contacts in general, may push users out.

Diffusion research separates simple from complex contagion (Centola and Macy, 2007). In simple contagion, one exposure can be enough. In complex contagion, people need reinforcement from several sources. Wide bridges of strong ties then matter more than long weak ties (Centola and Macy, 2007). Dense local ties alone do not guarantee fast spread, though (Manzo et al, 2018). Whether people joined Facebook depended on how many separate groups their contacts formed, not on how many contacts they had (Ugander et al, 2012). Weak ties carry information across sub-groups and strong ties within them (Friedkin, 1982). Users who bridge communities see and spread more information (Li et al, 2015). The best spreaders sit in the k-core, not among the most connected users (Kitsak et al, 2010). For these reasons we compare several mechanisms, a threshold rule among them, instead of assuming one.

Classic diffusion models formalize how a behavior passes along ties. In the independent-cascade model, each newly active contact gets one chance to activate a neighbor. In the linear-threshold model, a node activates once the share of its active neighbors passes a private threshold (Kempe et al, 2003). Diffusion models of this kind have also been used to simulate how harmful behavior such as aggression spreads on social media and to test strategies that limit it (Poiitis et al, 2022). We treat departure as the behavior and include both models among the mechanisms SILCOM compares.

Identification is the hard part. Correlated departures can come from contagion, from hidden homophily or from shared causes (Shalizi and Thomas, 2011; O'Malley et al, 2014). Anagnostopoulos et al (2008) proposed two tests for influence in social networks. The shuffle test permutes the times at which users adopt a behavior; correlation that survives shuffling points away from influence. The edge-reversal test reverses the direction of ties; influence should flow one way. Matched-sample estimation compares exposed and unexposed users with similar characteristics (Aral et al, 2009). Churn models built on such correlations can mix these up (Zhang et al, 2022). Network-wide

shocks, such as a migration after a blockchain hard fork, shift many users at once (Dileo and Zignani, 2025a,b). Proposed fixes include instrumental variables (O'Malley et al, 2014) and counterfactual data augmentation (Zhang et al, 2022). Homophily and contagion are, however, generically confounded in observational network data (Shalizi and Thomas, 2011). We pair a permutation test that keeps the timing of departures with a comparison of mechanisms on held-out data, and we score both against the tests above on synthetic data with known truth (Section 5). The result shows consistency with a mechanism. It does not prove cause.

2.3 Network robustness, targeted removal and targeting

Network science asks how networks break when nodes are removed at random or on purpose. Scale-free networks are thought to withstand random failure but to break when hubs are removed. Degree distributions can be optimized to resist waves of targeted and random attacks (Tanizawa et al, 2005). Robustness to random and to targeted removal trade off, and core–periphery structures balance the two best (Priester et al, 2014). Attacks based on centrality come close to the worst case (Trajanovski et al, 2013). Adding or rewiring a few edges can make a network much more robust (Beygelzimer et al, 2005).

We take two things from this work. The first is its outcome measures, such as the size of the largest component and the number of components. The second is the idea that new ties can protect a network, which motivates our reconnection intervention. But this work treats removal as purely structural. Removed nodes vanish, and the others do not react. In social systems, one departure changes the incentives of those who stay (Garcia et al, 2013). Hubs can sway whether their neighbors migrate (Galdeman et al, 2023). SILCOM therefore couples structural removal with behavioral rules, so a targeted shock can set off further departures.

The same centrality measures guide the opposite task: choosing whom to reach. Degree, betweenness (Freeman, 1977), PageRank (Brin and Page, 1998), the k-shell (Kitsak et al, 2010) and collective influence (Morone and Makse, 2015) each rank nodes by their structural importance for spreading. Influence maximization picks seed sets under cascade models (Kempe et al, 2003). Retention is a different goal, however. The users worth reaching are those likely to leave, and structural importance need not mark them. We test these rankings as targeting rules for support (Section 5).

2.4 Churn prediction

Churn prediction is well developed in marketing and online services. Hazard and survival models estimate when churn is likely, with time-varying covariates and user differences (Jamal and Bucklin, 2006; Kimitei et al, 2025). Social features help. Proportional-hazards models (Cox, 1972) remain a standard baseline. Collective classification, which uses links between users, beat standard classifiers on a mobile social network (Oentaryo et al, 2012). Graph convolution propagates user features along ties; its simplified form performs close to deeper graph networks at a fraction of the cost (Wu et al, 2019). In online games, graph neural networks (Han et al, 2024), dynamic graphs (Lee and Woo, 2025) and multiplex embeddings (Guo and Liu, 2026) improved

prediction. A contagion-aware graph model with a network risk index improved both prediction and the targeting of retention efforts (Lu et al, 2026). Counterfactual modeling has been used to separate a user's own intention to leave from social influence (Zhang et al, 2022).

These methods serve platform operators who hold full, private data on one centralized service. They predict who will leave. They do not simulate what happens to a community after someone leaves, and they cannot test interventions in advance. SILCOM uses hazard models as one part. It adds a prediction module checked against baselines, and it feeds its scores into a simulation where interventions can be tested.

2.5 Decentralized social networks

Architecture and measurement

Bluesky runs on the AT Protocol, which lets third parties provide each part of the platform (Kleppmann et al, 2024; Balduf et al, 2024). Each account's records sit in a signed public repository on a personal data server. Its networks have heavy-tailed distributions, high clustering and short paths (Quelle and Bovet, 2025). Opening to the public in February 2024 brought a surge of new users and little toxicity (Sahneh et al, 2025; Nogara et al, 2026). Scientists moved to it in late 2024 (Arroyo-Machado et al, 2025). Its public moderation data show widespread community blocking (Sokoto et al, 2026). In practice, Bluesky remains partly centralized (Yu et al, 2025). Mastodon, the other large decentralized platform, is a federation of independently run servers, called instances, each with its own topics and languages (Zignani et al, 2018). Early measurements found its follow network different from commercial microblogging sites (Zignani et al, 2018). Later work studied ties within and across instances and found boundary-spanning roles (La Cava et al, 2022). Others examined polarization between instances (La Cava et al, 2024) and echo chambers (Huitema et al, 2025). User activity on Mastodon has been compared with X (Hironaka et al, 2024). Collecting Mastodon data is hard because data sit on many instances with different access rules (Gao et al, 2025).

Governance

These platforms hand governance to communities. Mastodon instances set their own rules, which focus on harassment and hate (Nicholson et al, 2023). Rules grow longer as instances grow (Muralidharan et al, 2026). Shared blocklists protect communities but can also cut users off from support (Melder et al, 2025). Researchers have called for tools that help communities govern together (Hwang et al, 2026). Struett et al (2024) list governance, moderation at scale and commercial capture among the threats to the fediverse.

Migration

Users with more followers on X were more likely to migrate (Radivojevic et al, 2025). Migration patterns and the obstacles to staying differed across Mastodon, Bluesky and Threads (Jeong et al, 2024). Journalists seldom left Twitter fully (Ng and Ray, 2025). Misinformation pushed them away, and community values drew them to Mastodon.

Migration is both a personal and a collective act (Ittefaq, 2025; Tkacz et al, 2026). Moves to Mastodon followed a pattern of social influence (La Cava et al, 2023). Influence spread faster in sparser communities, where members signaled commitment and stressed shared identity. On blockchain-based platforms, hubs swayed whether their neighbors migrated (Galdeman et al, 2023). Migration also changed user behavior in ways visible in temporal graph models (Dileo and Zignani, 2025b).

These studies measure decentralized platforms at one point in time or around a migration. As far as we know, none models how users gradually leave communities already formed there, or tests responses to it.

2.6 Data-driven agent-based modeling

Agent-based models show how individual choices and network structure add up to collective outcomes. Their value depends on how well they are grounded in data. Social-science ABMs need empirical data both to calibrate and to validate them (Boero and Squazzoni, 2005). Large-scale behavioral data can inform agent rules directly (Kavak et al, 2018). Likelihood-based Bayesian estimation can beat simulated minimum-distance methods, except for large models with many parameters (Platt, 2020). Validation can also work at the meso level, by clustering simulated and real behavior (Bektas et al, 2021). Online, ABMs have been calibrated from retweet networks (Mastroeni et al, 2020). A Reddit dataset was built to ground ABMs in data (Sittar et al, 2026). Calibrated simulations have explained why diffusion differs between communities (Manzo et al, 2018). On decentralized platforms, an ABM using Steemit and Pokec data compared how misinformation spreads and tested governance tools against it (Malgonde et al, 2025).

Evidence on real interventions is mixed. On Stack Exchange, a badge marking newcomers reduced unwelcoming replies for a while but did not stop retention from falling (Santos et al, 2020). Offline meetups did raise Wikipedia contributions (Schwitter, 2024). Testing interventions in a calibrated simulation first can save effort and avoid harm. The simulation must be validated, though, and its results treated as hypotheses.

2.7 Summary and research gap

No framework yet joins these streams. Studying withdrawal cascades on a decentralized platform takes several steps. One must collect scattered public data and rebuild histories from it. One must test whether departures cluster beyond chance and compare mechanisms on held-out data. One must compare each step with established methods. Finally, one must test interventions on the calibrated model. Silcom provides these steps as one reusable, open system.

3 The Silcom framework

3.1 Design goals

Four goals follow from the gaps above:

Table 1 What each research stream offers and what it leaves open for studying withdrawal cascades in decentralized networks.

Stream	Offers	Leaves open
Community sustainability (Kairam et al, 2012; Ribeiro, 2014; Walk et al, 2016; Garcia et al, 2013)	Predictors of community survival; activity dynamics; cost–benefit cascade model	How decline unfolds over time; behavioral response to departures; decentralized platforms
Social contagion (Wu et al, 2013; Kempe et al, 2003; Anagnostopoulos et al, 2008; Shalizi and Thomas, 2011)	Evidence that departures cluster; diffusion models; tests for influence	Which test works for departures; comparing mechanisms on held-out data
Robustness and targeting (Trajanovski et al, 2013; Kitsak et al, 2010; Morone and Makse, 2015)	Targeted vs. random removal; fragmentation metrics; centrality rankings	Users' reactions to the removal of others; targeting for retention
Churn prediction (Oentaryo et al, 2012; Zhang et al, 2022; Lu et al, 2026)	Accurate individual risk scores using social features	Public data; simulation of interventions; decentralized settings
Decentralized networks (Kleppmann et al, 2024; Quelle and Bovet, 2025; La Cava et al, 2023; Struett et al, 2024)	Structure, activity, governance and migration of Bluesky and Mastodon	Gradual withdrawal after migration; testing responses to it
Data-driven ABMs (Boero and Squazzoni, 2005; Platt, 2020; Sittar et al, 2026)	Calibration and validation methodology	Application to withdrawal on decentralized platforms; reusable tools

G1 One-time collection. The framework works from a single collection of public data and rebuilds history after the fact. It can be rerun on any community without months of monitoring.
G2 Decentralized data. It reads data spread across personal data servers run by different providers, through the open protocol rather than a platform API.
G3 Validation. Every model is checked against held-out real data and against published alternatives. Where real data offer no ground truth, it is checked against a synthetic benchmark.
G4 Reuse. Modules are separate, configurable and open, so others can swap a part or study a new community.

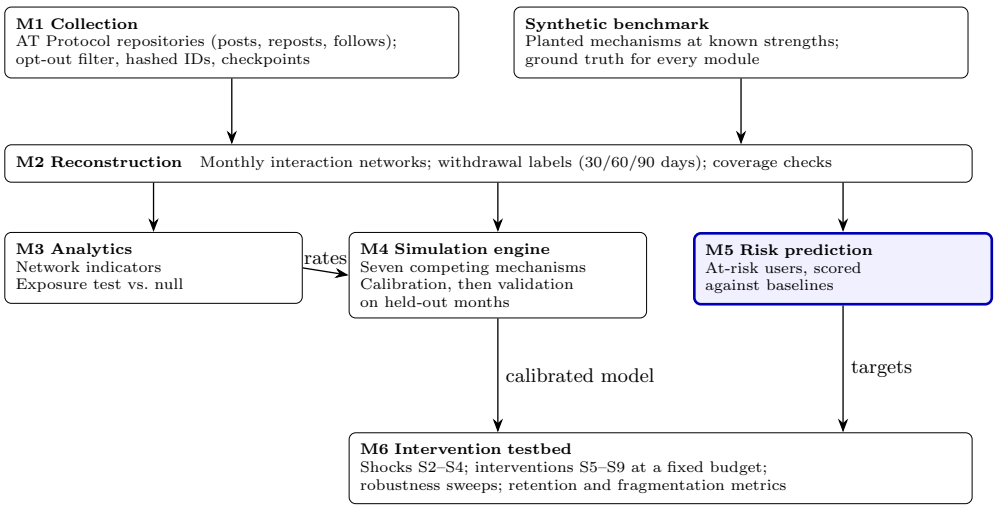

Fig. 1 Architecture of Silcom. Real and synthetic data share one reconstruction step. The highlighted risk prediction module (M5) also supplies targets for the intervention testbed (M6).

3.2 Architecture overview

Figure 1 shows the six modules. Real data from collection (M1) and synthetic data from the benchmark pass through the same reconstruction module (M2). Every later module can thus be tested on known ground truth before it meets real data. Analytics (M3) computes network indicators, runs the exposure test and supplies baseline withdrawal rates to the simulation engine (M4). Risk prediction (M5) scores users. The intervention testbed (M6) combines the calibrated simulation with these scores. All modules share one event schema: hashed source, hashed target, interaction type and timestamp. Parts can therefore be replaced one at a time. Collection runs once; every other module runs offline from the saved data.

3.3 M1: Data collection

Sampling

Silcom samples by breadth-first snowball sampling from seed accounts. It reads up to three pages of each account's follow list (100 accounts per page) through Bluesky's public AppView. Unseen accounts join the queue until the sample reaches a target size. Our sample began from 12 seed accounts, chosen as active members spread across sub-groups of one community, and reached 3,129 accounts. Snowball samples favor well-connected users; we return to this in Section 11.

Repositories

Each Bluesky account keeps a public repository of signed records on a personal data server (Kleppmann et al, 2024). Silcom finds each account's server through the PLC directory, or through the web domain for `did:web` identifiers. It then pages through the post, repost and follow collections with `com.atproto.repo.listRecords`, newest

first. Paging stops after 50 records in a row older than the window. A post counts as a reply if it has a parent reference, as a quote if it embeds a record, and as an original post otherwise. Mentions come from its rich-text facets. Records dated after the collection date are dropped. Because records are read from the repository rather than from a feed, the history can be rebuilt after the fact, as G1 requires.

Throughput and robustness

Requests are rate-limited and back off on HTTP 429 and server errors. The User-Agent header names the study. Events are saved in chunks along with the list of finished accounts, so an interrupted run resumes without repeating work. On a free Google Colab instance, 3,000 repositories took 1 hour 52 minutes.

Privacy

Each account identifier is replaced at collection time by a salted SHA-256 hash, cut to 16 hexadecimal characters. The salt stays private. No post text, display name or profile data is stored. Accounts carrying the `!no-unauthenticated` label, which asks clients not to show them to logged-out viewers, are skipped.

3.4 M2: Reconstruction, cleaning and withdrawal labeling

Cleaning

Exact duplicate records are removed. Some accounts share at least 20 identical activity timestamps with another account, for example when one scheduling tool posts to both. These are treated as linked or automated and excluded.

Observation window

Let T_{end} be the collection date. The window covers the 18 months before it, indexed $m = 0, \ldots, 17$. With collection on 1 November 2026, it runs from 1 May 2025 to 1 November 2026. Months 0–11 are for calibration and later months for validation. Self-interactions and events outside the window are dropped.

Activity and withdrawal

A user's own activity is posts, replies, quotes and reposts. For user i, let f_i and ℓ_i be the first and last activity in the window, and n_i the number of actions. User i is *eligible* if

$$n_i \geq 5 \quad \text{and} \quad T_{\text{end}} - f_i \geq \tau + 30 \text{ days}, \tag{1}$$

where τ is the inactivity threshold. An eligible user has *withdrawn* if $T_{\text{end}} - \ell_i \geq \tau$. The withdrawal month is then $w_i = \text{month}(\ell_i)$; otherwise $w_i = \infty$ and the user is censored. We use $\tau = 60$ days and test 30 and 90 days. A withdrawal can only be confirmed τ days after the last activity. Months after $m^* = \text{month}(T_{\text{end}} - \tau) - 1 = 15$ are therefore left out of any outcome that needs a confirmed withdrawal.

Interaction networks

For months $[a, b)$, SILCOM builds a directed, weighted network $W_{[a,b)}$ among eligible users. W_{ij} counts the replies, quotes, reposts and mentions from i to j. Its undirected, unweighted form A ($A_{ij} = 1$ if $W_{ij} + W_{ji} > 0$) defines neighborhoods. We use interactions instead of follows because they carry timestamps and reflect active ties.

Coverage diagnostics

SILCOM counts sampled accounts with no activity of their own in the window. It splits them into accounts for which the server returned other records, such as follows, and accounts with no records at all. Follows are kept only inside the window, so "no records" is an upper bound on failed collections.

Robustness checks

Network measures depend on sample size. SILCOM therefore also builds a subsample of the first 1,350 accounts in collection order, reached by the same snowball process. It also recomputes network summaries without quote posts.

3.5 M3: Network analytics and the exposure test

Indicators

For each eligible user, SILCOM computes in- and out-strength and degree. It computes reciprocity, the share of a user's partners who interact back. Partner diversity is the normalized entropy of a user's outgoing weights. It also computes local clustering, k-core number (Seidman, 1983) and approximate betweenness from 300 sampled sources. Communities come from the Louvain method on the weighted undirected network. At network level, it reports mean partners, density, reciprocity, mean clustering, the largest component's share of users, modularity, the top 1%'s share of interaction volume and the Gini coefficient of strength. Time to withdrawal is summarized with Kaplan–Meier curves.

Exposure test

Do departures cluster in the network beyond chance? SILCOM answers with a discrete-time hazard model on user-months. In month $m \le m^*$, users who joined before m and have not yet left are at risk. Exposure is measured three ways. The first is the share of a user's neighbors who have already left,

$$E_i(m) = \frac{\sum_j A_{ij} \mathbb{1}[w_j < m]}{\sum_j A_{ij}}. \tag{2}$$

The second is the log of one plus their number, and the third flags whether any neighbor has left. The model is

$$\text{logit}\ \Pr(w_i = m \mid w_i \ge m) = \alpha_m + \beta\, x_i(m) + \gamma_1 \log(1 + \bar{a}_i(m)) + \gamma_2 \log(1 + k_i). \tag{3}$$

Here $x_i(m)$ is one exposure measure and α_m are month fixed effects, which absorb platform-wide change. $\bar{a}_i(m)$ is mean monthly activity over the past three months, and k_i is degree. Network position could yield $\beta > 0$ without any social effect. We therefore compare $\hat{\beta}$ with a null in which users' network positions are shuffled while their own activity and departure times stay fixed. The p-value is $(1 + \#\{\hat{\beta}_\pi \geq \hat{\beta}\})/(N_{\text{perm}} + 1)$, with 200 permutations (100 for the subsample). If an estimate's 95% interval spans more than a factor of 100, we call it undetermined. A significant result shows that departures cluster beyond what timing, activity and degree explain. It does not, by itself, prove contagion (Shalizi and Thomas, 2011; O'Malley et al, 2014); Section 5 tests how far it can be trusted.

3.6 M4: Simulation engine

Agents and network

A simulation starts at month s. Agents are eligible users who were active in the two months before s and had not left. They are linked by the interactions of the previous four months. Each agent has an activity rate r_i (mean monthly actions over the past three months), a degree k_i and a community.

Withdrawal rule

In each simulated month t, an active agent leaves with probability $h_i(t)$:

$$\operatorname{logit} h_i^0(t) = \underbrace{\hat{\gamma}_0 + \hat{\gamma}_1 \log(1 + r_i) + \hat{\gamma}_2 \log(1 + k_i)}_{\eta_i} + \delta + \phi_i(t) - u_i(t). \tag{4}$$

The baseline η_i comes from a hazard model fitted on the calibration months, without exposure or month effects. δ is a calibrated offset. $u_i(t)$ is the effect of an intervention (Section 3.8). $\phi_i(t)$ and the rules below define the mechanism being tested. SILCOM compares seven mechanisms, five of its own and two classic diffusion models:

- **Independent:** $h_i = h_i^0$ with $\phi_i = 0$. Users leave at their own rate.
- **Influence:** $\phi_i(t) = \beta E_i(t)$, the share of neighbors who have left. This is the fraction rule observed by Wu et al (2013).
- **Reciprocity loss:** $\phi_i(t) = \beta L_i(t)$. Here $L_i(t) = 1 - \sum_j W_{ji}\, a_j(t) / \sum_j W_{ji}$ is the share of interactions once received that were lost because their senders left; $a_j(t) \in \{0, 1\}$ marks whether j is active.
- **Combined:** $\phi_i(t) = \beta\,(E_i(t) + L_i(t))$.
- **Threshold:** the agent also leaves for certain once its active neighbors drop below $\theta k_i(0)$, a k-core-style cost–benefit rule in the spirit of Garcia et al (2013).
- **Independent cascade** (Kempe et al, 2003): each neighbor who left in the previous month gets one independent chance p to make the agent leave, so $h_i = 1 - (1 - h_i^0)(1 - p)^{n_i(t)}$, where $n_i(t)$ counts neighbors who left in month $t - 1$.
- **Linear threshold** (Kempe et al, 2003): each agent draws a private threshold $\theta_i \sim U(0, c)$ once; it also leaves once $E_i(t) \geq \theta_i$.

Calibration

Each mechanism is calibrated by grid search on months 4–11. The loss adds the fit of the retention curve to the fit of departure clustering:

$$\mathcal{L} = \sqrt{\tfrac{1}{T}\textstyle\sum_t \left(S_t^{\text{sim}} - S_t^{\text{obs}}\right)^2} + \lambda \left|\rho^{\text{sim}} - \rho^{\text{obs}}\right|, \qquad \lambda = 0.05. \tag{5}$$

S_t is the share of agents still active. ρ is withdrawal assortativity: the mean share of departed neighbors among departed agents, divided by the overall departure share. The grid spans $\delta \in \{-1.5, -1.25, \ldots, 0.5\}$; $\beta \in \{0.1, 0.25, 0.5, 1, 2, 3, 4\}$ (up to 3 for the combined mechanism); $\theta \in \{0.02, 0.05, 0.1, 0.2, 0.3\}$; $p \in \{0.0025, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2\}$; and $c \in \{2, 4, 8, 16, 32\}$. Each grid point averages 10 runs.

Validation and selection

Each calibrated mechanism then runs, unchanged, from month 12 to m^* over 50 runs and is scored by the same loss. Mechanisms within 1% of the best validation loss count as tied. Ties go to the simplest mechanism, in the order independent, influence, reciprocity loss, threshold, independent cascade, linear threshold, combined. The independent model is the baseline: a social mechanism wins only if it beats a model with no social effects on held-out data (Boero and Squazzoni, 2005). Two checks test the selection. First, we rank mechanisms under clustering weights $\lambda \in \{0, 0.02, 0.04, 0.05, 0.1, 0.2\}$. Second, we calibrate on months 4–8 and validate on months 9–11, with the baseline hazard refitted on months before 9.

3.7 M5: Risk prediction

For each at-risk user and prediction month m, the label is whether the user leaves within three months. Features use only data from before m. They fall into three groups. *Activity* covers actions in the last month and the last three months, trend, months since last activity and tenure. *Network* covers degree, in- and out-strength, reciprocity, partner diversity, clustering and k-core number, all from the previous four months. *Exposure* covers $E_i(m)$, the number of departed neighbors, the drop in interactions received, and the community's departure share. Models train on prediction months 4–9, whose labels fall wholly in the calibration period, and are tested on months 12–13. The model used downstream is chosen on the last two training months, so test months never inform a choice. Candidates are logistic regression on activity alone, logistic regression on all features, a random forest and gradient boosting. We report ROC AUC with 95% bootstrap intervals (200 resamples), precision–recall AUC and precision among the riskiest 5%. An ablation drops one feature group at a time.

3.8 M6: Intervention testbed

Setups

The testbed starts at month 12 and runs for 12 months on the 2,106 agents active then. It runs two setups: the selected mechanism and the independent mechanism. Comparing them shows which conclusions depend on social effects.

Table 2 Simulation scenarios.

Scenario	Shock at $t = 0$	Intervention (budget 2% of agents per month)
S1 Baseline	None	None
S2 Random withdrawal	5% at random	None
S3 Hub withdrawal	5% highest degree	None
S4 Supportive-user withdrawal	5% highest reciprocated volume	None
S5 Reconnection	As S3	Each exposure-targeted agent gains one tie to a random active member of its community
S6 Exposure-targeted support	As S3	Exposure-targeted agents' withdrawal log-odds reduced by $u = 0.5$ for that month
S7 Adaptive	As S3	S5 and S6 combined, restricted to communities in the top quintile of members lost
S8 Risk-targeted support	As S3	Support ($u = 0.5$) for agents with the highest M5 predicted risk
S9 Oracle-targeted support	As S3	Support ($u = 0.5$) for agents with the highest simulated hazard (upper bound)

Scenarios

Table 2 lists nine scenarios. Shocks remove 5% of agents at $t = 0$, chosen at random, by highest initial degree (hubs), or by highest reciprocated volume $\sum_j W_{ij} \mathbb{1}[W_{ji} > 0]$ (supportive users). Interventions act after the hub shock. Exposure-targeted interventions (S5–S7) reach at-risk agents, $\max(E_i(t), L_i(t)) > 0.3$, in order of risk. S8 targets the agents with the highest M5 risk at the start. S9 targets those with the highest current simulated hazard. No real intervention could do this, so S9 is an upper bound. All share a budget of 2% of agents per month.

Paired cascades

A naive cascade measure compares departures after a shock with departures without it. This is biased. Random shocks often remove low-activity users who would have left anyway, so random removal seems to *prevent* departures. SILCOM uses common random numbers instead. For each seed, the no-shock run (S1) and the shock run share the same random draws. The cascade counts extra departures among agents *not* removed by the shock:

$$C = \sum_{i \notin \mathcal{S}} \mathbb{1}[i \text{ leaves} \mid \text{shock}] - \sum_{i \notin \mathcal{S}} \mathbb{1}[i \text{ leaves} \mid \text{no shock}], \tag{6}$$

where $\mathcal{S}$ is the set of removed agents. Under the independent mechanism, $C = 0$ exactly; the framework checks this on every run. Each scenario runs over 30 seeds, and scenarios are compared seed by seed with Wilcoxon signed-rank tests.

Outcomes

Each run records retention, the largest component's size relative to the start, the number of components, the reciprocity of the remaining network, and the paired cascade, also as a percentage of the shock. Cost is ties added plus supported agent-months. Under the selected mechanism, one-at-a-time sweeps vary social strength, shock size, budget, support effect and risk threshold over 15 seeds (Appendix B).

3.9 Synthetic benchmark and system tests

Benchmark

The generator draws a power-law cluster network. Each user gets a log-normal activity rate and a start month. Users then leave month by month under a planted mechanism: independent, influence ($\beta = 3$), reciprocity loss ($\beta = 3$), threshold ($\theta = 0.5$), independent cascade ($p = 0.1$) or linear threshold ($c = 2$). Only active users interact. The output follows the M1 schema and runs through the whole pipeline: reconstruction, the exposure test (50 permutations), calibration and selection (a smaller grid with 4 calibration and 10 validation runs). We make five replicates of 1,500 users per mechanism and record the selected mechanism and the exposure test result.

Labeling check

Labeling withdrawal by inactivity at the window's end may catch users who are only pausing. To measure this, we generate users who leave at a constant true rate of 1% or 2% per month, with no social effects, in four replicates of 3,000 users. We compare the measured monthly rate in the calibration and validation periods.

Scalability

We time each module and track memory on synthetic networks of 1,000 to 50,000 users.

3.10 Implementation

SILCOM is written in Python. It uses `requests` for collection and `pandas` for events. Networks and simulation use `networkx` and `scipy.sparse`. Statistics use `statsmodels`, and prediction uses `scikit-learn`. Simulations run on sparse matrices, so each month's update is a few matrix–vector products. All parameters are set in one configuration block (Appendix A). This study uses only publicly available data from Bluesky. Code, the benchmark, configuration files and pseudonymized data are available on OSF (https://osf.io/6kdfj, view-only link for review).

Table 3 Sample, coverage, withdrawal and network structure. Withdrawal is the share of eligible users inactive for at least the given number of days before the collection date. Network measures use interactions among eligible users over the whole window.

	Full sample	First 1,350 accounts
Sampled accounts	3,121	1,350
Active in window	2,804	1,290
No activity, other records returned	135	–
No records at all	182 (5.8%)	–
Eligible users	2,522	1,152
Withdrawn, 30 days	25.3%	24.7%
Withdrawn, 60 days (main)	19.3%	19.2%
Withdrawn, 90 days	16.0%	16.1%
Interaction edges	65,105	15,513
Mean partners	25.8	13.5
Reciprocity	0.32	0.36
Mean clustering	0.32	0.31
Largest component share	0.94	0.92
Modularity (Louvain)	0.54	0.58
Top 1% share of interaction	0.38	0.39
Strength Gini	0.83	0.81

4 Evaluation of the framework

4.1 Reconstruction fidelity (RQ1)

Data and cleaning

We collected 3,129 Bluesky accounts. Cleaning removed 1,493 exact duplicate records (0.012% of records). It also dropped eight accounts that shared at least 20 identical timestamps with another account. That left 3,121 accounts, of which 2,522 were eligible (Table 3).

Coverage

317 accounts posted nothing in the window. For 135 of them the server returned other records, so collection worked. 182 accounts (5.8%) returned no records at all. Follows are kept only inside the window, so this is an upper bound on failures. With only hashed IDs, an account dormant all window looks just like a failure. Either way, accounts without records never enter the analysis and cannot inflate departure rates.

Withdrawal labels

Departure rates fall as the inactivity threshold rises, from 25.3% at 30 days to 16.0% at 90 days. Going from 30 to 60 days removes about a quarter of departures, which suggests that many users pause. The subsample of the first 1,350 accounts had the same rate as the full sample (19.2% vs. 19.3%).

Network

The interaction network was well connected (94% of users in the largest component), clustered (0.32) and modular (0.54). Interaction was concentrated. The top 1% of users took part in 38% of all interaction, and the Gini coefficient of strength was 0.83. Removing quote posts changed none of these by more than 0.02 (Appendix C).

A labeling bias at the window's end

Measured departures rose steeply over the window. Among simulation agents, the monthly rate averaged 0.79% in calibration months and 2.36% in validation months, three times higher. Part of this rise comes from the definition itself. With a constant true rate of 2% per month, the measured rate was 1.27 times higher in validation months. At 1% per month it was 1.59 times higher. A user who posts rarely can easily show a 60-day gap at the end of the window. The same gap earlier is invisible, because the user posts again afterwards. The bias explains part of the rise, not all. The rest may be a real increase in departures. Any study that labels departure by inactivity faces this bias. It also affects our simulations, because no mechanism fitted on earlier months can foresee it.

4.2 Mechanism identification (RQ2)

Benchmark: detection works, naming does not

We ran five replicates for each of six planted mechanisms (Figure 2). The exposure test never flagged clustering when none was planted (0 of 5 at $p < 0.05$). It caught clustering in all 25 replicates with a social mechanism. Simulation-based selection did far worse. It picked the planted mechanism in 7 of 30 replicates (23%). It told social from independent departures apart in 25 of 30 (83%). It recovered the threshold rule in 4 of 5 replicates. The linear-threshold model was taken for the threshold rule in all 5, and reciprocity loss in 3 of 5. The independent cascade was recovered once. With no social mechanism planted, selection still chose a social one in 3 of 5. The exposure test is thus the framework's reliable evidence of social withdrawal. The simulation can add weight to "social over independent". It cannot say which social mechanism is at work.

Exposure test on real data

The share of a user's contacts who had left predicted the user's own departure the next month (Table 4). The odds ratio was 3.75 per unit exposure (95% CI 1.77–7.94), with activity, degree and month controlled. That is about 14% higher odds for each extra 10 points of departed contacts. The estimate beat all 200 permutations ($p = 0.005$, the smallest possible value), whose mean odds ratio was 0.96. The other two measures agreed. Each extra log-unit of departed contacts raised the odds by 59%, and having any departed contact raised them by 54% (both $p = 0.005$).

The subsample of 1,350 accounts did not reproduce the result. With 215 departures, its share-based odds ratio was 1.97 (95% CI 0.59–6.56; permutation $p = 0.089$). The count and any-contact measures showed nothing. Clustering is solid in the full sample, but detecting it took about 2,500 eligible users and 475 departures. Studies of smaller communities should expect low power.

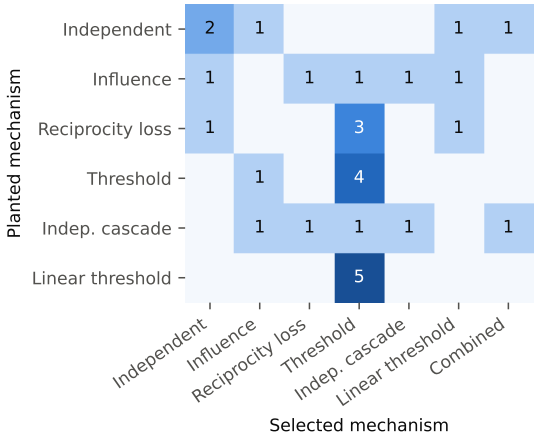

Fig. 2 Mechanism selection on synthetic data with a planted mechanism (five replicates of 1,500 users each). Rows sum to five. The exposure test, not shown, detected every planted social mechanism and raised no false alarm.

Table 4 Exposure test: discrete-time hazard of withdrawal on exposure to departed contacts, with month fixed effects and controls for activity and degree. Odds ratios with 95% CI; permutation p from shuffling network positions (200 permutations; 100 for the subsample).

Sample	Exposure measure	User-months	Departures	OR (95% CI)	Perm. p
Full	Share of contacts gone	34,035	475	3.75 (1.77–7.94)	0.005
	log(1 + contacts gone)			1.59 (1.13–2.23)	0.005
	Any contact gone			1.54 (1.14–2.10)	0.005
Subsample	Share of contacts gone	15,579	215	1.97 (0.59–6.56)	0.089
	log(1 + contacts gone)			0.89 (0.48–1.67)	0.723
	Any contact gone			0.90 (0.53–1.53)	0.683

Mechanism selection on real data

We had 2,172 calibration and 2,106 validation agents. The independent-cascade model had the lowest validation loss, 10% below the next model (Table 5). Every social mechanism except the threshold rule fitted the calibration months two to four times better than the independent model (loss 0.003–0.006 vs. 0.013). All predicted the same retention curve on validation months (error 0.038–0.039). They differed in how well they reproduced the clustering of departures. Observed withdrawal assortativity was 0.58. The independent cascade reached 0.36; the other probabilistic mechanisms reached 0.22–0.26. The threshold rule matched retention best (error 0.033), but only through a higher baseline. It fitted the calibration months poorly (loss 0.038) and produced far too much clustering (1.31).

The independent cascade won under every clustering weight from 0.02 to 0.2. With no weight on clustering, the threshold rule won through its retention fit alone. On the earlier split (calibrate on months 4–8, validate on 9–11), reciprocity loss won and the independent cascade ranked fifth. The best social mechanism beat the independent

Table 5 Mechanism selection on Bluesky. Each mechanism is calibrated on months 4–11 and validated, unchanged, on months 12–15. Loss = retention RMSE + 0.05 × |clustering gap|. Clustering: simulated withdrawal assortativity; observed on validation months: 0.58. "Edge": best value at the edge of the grid. The k-core row comes from the comparison in Section 5.2. Sources: [a]Kempe et al (2003); [b]Wu et al (2013); [c]Garcia et al (2013).

Mechanism	Cal. loss	Val. RMSE	Clustering	Val. loss	Fitted parameters
Independent cascade[a]	0.0052	0.0384	0.36	**0.0494**	$p = 0.005$, $\delta = -0.5$
Linear threshold[a]	0.0042	0.0383	0.25	0.0547	$c = 32$ (edge), $\delta = -0.5$
Combined (Silcom)	0.0032	0.0386	0.26	0.0548	$\beta = 0.5$, $\delta = -0.5$
Reciprocity loss (Silcom)	0.0056	0.0386	0.24	0.0555	$\beta = 1$, $\delta = -0.5$
Share of contacts gone[b]	0.0054	0.0386	0.24	0.0558	$\beta = 0.5$, $\delta = -0.5$
Independent (baseline)	0.0125	0.0386	0.22	0.0568	$\delta = -0.5$
Threshold (Silcom)	0.0379	0.0328	1.31	0.0690	$\theta = 0.02$ (edge), $\delta = -0.25$
k-core collapse[c]	0.0379	0.0328	1.31	0.0690	$k = 1$, $\delta = -0.25$

Table 6 Risk prediction (M5): withdrawal within three months, trained on prediction months 4–9 and tested on months 12–13 (base rate 8.6%). ROC AUC with 95% bootstrap CI. The model selected on held-out training months is marked.

Model	ROC AUC (95% CI)	PR AUC	Precision, top 5%
Logistic, activity only	0.831 (0.811–0.847)	0.286	0.342
Logistic, all features (selected)	0.827 (0.806–0.844)	0.288	0.360
Random forest	0.829 (0.806–0.845)	0.288	0.360
Gradient boosting	0.823 (0.804–0.839)	0.251	0.284
Selected model without activity	0.773	0.208	0.182
Selected model without network	0.832	0.289	0.356
Selected model without exposure	0.827	0.290	0.356

model under every weight and on both splits. As in the benchmark, we conclude that Bluesky departures fit a social mechanism better than an independent one. The independent cascade is the best current description, not a proven mechanism.

4.3 Risk prediction (RQ3)

Departure within three months was predictable (Table 6). Training used 13,957 user-months with 391 departures. The test months held 4,497 user-months and 385 departures, a base rate of 8.6%. Logistic regression on activity alone reached an AUC of 0.83. Of the riskiest 5%, 34% left, four times the base rate. The model chosen on held-out training months used all features (AUC 0.83; 36% precision among the riskiest 5%). Network and exposure features added little. Dropping activity features cut the AUC to 0.77. Dropping network or exposure features changed it by less than 0.01. At community level, mean predicted risk rose with actual departures, but not significantly ($\rho = 0.45$, 13 communities, $p = 0.12$).

So exposure to departed contacts is linked to leaving (Section 4.2), yet adds little to prediction once own activity is known. Both can be true. The exposure effect is real but small next to the spread in activity, and part of it runs through activity.

Table 7 Summary of the comparisons. Each SILCOM module was run against published methods for the same task, on the same data.

Task	Compared methods	Outcome
Detecting clustered departures (M3)	Hazard-model p-value; shuffle and edge-reversal tests (Anagnostopoulos et al, 2008)	All but edge reversal detect planted contagion; none is immune to homophily; on Bluesky, the permutation and shuffle tests agree
Modeling cascades (M4)	Independent cascade, linear threshold (Kempe et al, 2003); share of contacts gone (Wu et al, 2013); k-core collapse (Garcia et al, 2013)	The independent cascade fits held-out months best and is adopted as SILCOM's selected mechanism
Predicting churn (M5)	Cox model (Cox, 1972); simplified graph convolution (Wu et al, 2019); k-core and exposure rules (Garcia et al, 2013; Wu et al, 2013)	M5 ties with the graph model (AUC 0.83) and beats the rest
Targeting support (M6)	Degree, betweenness (Freeman, 1977), k-shell (Kitsak et al, 2010), PageRank (Brin and Page, 1998), collective influence (Morone and Makse, 2015), random	Risk targeting keeps 3.5 times more users than random; centrality targeting keeps almost none

4.4 Scalability

Reconstruction, the exposure model and simulation grew roughly linearly with the number of users (Appendix C). At 50,000 users and 4.3 million events, reconstruction took 6 seconds, fitting the exposure model 2 seconds and one 12-month simulation run 10 seconds. Network analytics took longest, 170 seconds, mainly for betweenness and community detection. Memory stayed below 2.9 GB. The full analysis, including the benchmark and every comparison in Section 5, ran in under an hour on a free Colab instance.

5 Comparison with published methods (RQ4)

We compared each module with published methods for the same task, using the same data, splits and budgets. Table 7 summarizes the outcome; the subsections give the detail.

5.1 Detecting clustered departures

Methods

We compared SILCOM's permutation test with three alternatives. The first is the hazard model's own Wald p-value, without permutation. The second is the shuffle test of Anagnostopoulos et al (2008). It estimates the social-correlation coefficient α, the

Table 8 Tests for clustered departures. Left: share of synthetic replicates with $p < 0.05$ (five replicates per planted truth; contagion pools influence, reciprocity loss and threshold). Right: p-values on Bluesky. Shuffle and edge-reversal tests after Anagnostopoulos et al (2008).

	Synthetic: share significant			Bluesky
Test	No effect	Homophily	Contagion	p
SILCOM permutation test	0.0	0.4	1.0	0.005
Hazard model, no permutation	0.0	0.2	1.0	<0.001
Shuffle test	0.0	0.4	1.0	0.040
Edge-reversal test	0.0	0.0	0.33	0.79

effect of log(1+contacts gone) on the monthly hazard with month fixed effects. It then shuffles departure months among users who left and re-estimates α 200 times; influence should make the real α exceed the shuffled ones. The third is their edge-reversal test. It compares α computed over the users one interacts *with* against α over the users who interact with one, with a user-level bootstrap (100 resamples). Influence should flow along the first direction. All four were scored on synthetic data with five planted truths, five replicates each: no social effect, influence, reciprocity loss, threshold, and *homophily*. Under homophily, neighbors share a latent propensity to leave but do not affect one another.

Results

No test raised a false alarm when departures were independent (Table 8). The permutation test, the plain hazard model and the shuffle test detected every planted contagion. The edge-reversal test detected a third, mostly the threshold rule (4 of 5); our planted ties are symmetric, which gives it little to work with. Under homophily, however, the permutation and shuffle tests flagged clustering in 2 of 5 replicates and the plain model in 1 of 5. Only edge reversal stayed silent, at the cost of its low power. No test, then, separates contagion from homophily reliably, in line with theory (Shalizi and Thomas, 2011).

On Bluesky, the permutation test ($p = 0.005$) and the shuffle test ($p = 0.04$) agreed: departures cluster, and their timing matters. The shuffled coefficient was lower than the real one (-1.38 vs. -1.22; both are negative because this specification has no controls for degree, which lowers the hazard). The edge-reversal test found no direction ($p = 0.79$). We therefore report clustering beyond chance and do not claim contagion.

5.2 Modeling cascades

Methods

We compared SILCOM's own mechanisms with four published models of how departures spread: the independent-cascade and linear-threshold models (Kempe et al, 2003), the share-of-friends-gone rule of Wu et al (2013), and the k-core collapse model of Garcia et al (2013), in which a user leaves once fewer than k contacts remain ($k \in \{1, 2, 3, 5\}$). All went through the same calibration, validation and selection (Section 3.6).

Table 9 Churn prediction compared with published methods (withdrawal within three months; test months 12–13; 385 departures among 4,497 user-months). ΔAUC is M5 minus the method, with a paired-bootstrap p-value. Sources: [a]Wu et al (2019); [b]Cox (1972); [c]Garcia et al (2013); [d]Wu et al (2013).

Method	ROC AUC (95% CI)	PR AUC	Prec. top 5%	ΔAUC	p
SGC graph model[a]	0.829 (0.809–0.845)	0.282	0.369	−0.002	0.66
Silcom M5	**0.827 (0.806–0.844)**	**0.288**	0.360	–	–
Cox model[b]	0.810 (0.790–0.830)	0.224	0.240	0.017	<0.01
Degree heuristic	0.775 (0.754–0.794)	0.201	0.262	0.051	<0.01
k-core rule[c]	0.774 (0.753–0.794)	0.200	0.231	0.051	<0.01
Share of contacts gone[d]	0.459 (0.448–0.470)	0.091	0.058	0.366	<0.01

Results

Table 5 gives the result. The independent cascade fitted held-out months best, ahead of the linear threshold and Silcom's combined mechanism. Its advantage lay in clustering: it reproduced the observed tendency of departures to bunch among contacts better than any other probabilistic model. The k-core model settled at $k = 1$, where it behaves exactly like the threshold rule, and fitted worst. In a separate check, all models identified *which* agents left about equally well (agent-level AUC 0.76–0.78). Following the selection rule, we adopt the independent cascade as Silcom's mechanism for the scenarios in Section 6. Its lead depends on the validation months, as Section 4.2 showed, so we also report every scenario under the independent mechanism.

5.3 Predicting churn

Methods

On the same training and test months as M5, we compared the selected M5 model with four published approaches. A Cox proportional-hazards model (Cox, 1972) was fitted on monthly intervals with the M5 features. A simplified graph convolution (SGC) (Wu et al, 2019) propagated the standardized features twice over the normalized interaction network and fed the original and propagated features to a logistic regression. Two published churn rules score users by k-core number (Garcia et al, 2013) and by the share of contacts already gone (Wu et al, 2013). A degree heuristic completes the set. Differences in AUC were tested with a paired bootstrap (100 resamples).

Results

M5 matched the graph model: their AUCs differed by 0.002 ($p = 0.66$), and M5 had the higher PR AUC (Table 9). M5 beat the Cox model, the degree heuristic and the k-core rule by 0.02 to 0.05 in AUC (all $p < 0.01$). The share of contacts gone, used alone, ranked users worse than chance (AUC 0.46). One likely reason is that users with many departed contacts tend to be long-standing users who have outlasted them; in the hazard model, the same exposure raises risk only once activity and degree are held fixed. A simple model on own activity, then, does as well as a graph neural baseline here, and network rules on their own do poorly.

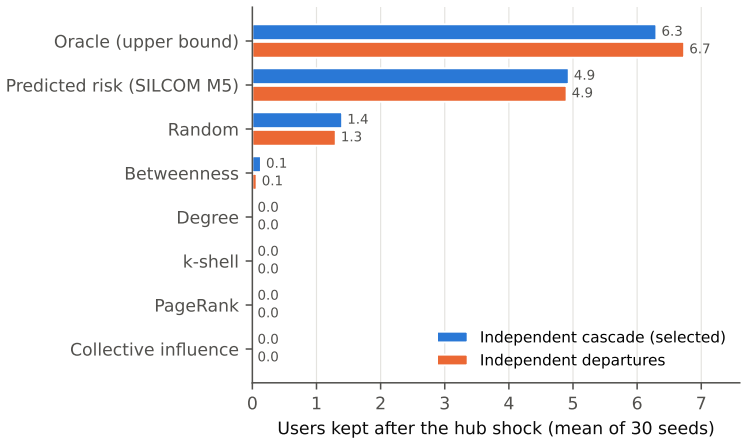

Fig. 3 Users kept by each way of targeting support, after the hub shock (2,106 agents, 12 months, budget 2% per month, 30 seeds). Risk targeting uses SILCOM's M5 scores; the oracle uses the simulation's own hazard.

5.4 Targeting support

Methods

M6 targets support by predicted risk (S8). We replaced that ranking with six others at the same budget (2% of agents per month, support effect $u = 0.5$): degree, betweenness (Freeman, 1977), k-shell (Kitsak et al, 2010), PageRank (Brin and Page, 1998), collective influence at radius 2 (Morone and Makse, 2015) and a random order. Each strategy ran after the hub shock over 30 seeds, paired with S3, under the selected mechanism and under the independent one. The oracle (S9) gives an upper bound.

Results

Risk targeting kept 4.9 users under both mechanisms, 78% of the oracle's gain under the independent cascade (Figure 3). Random targeting kept 1.4 and 1.3. Every centrality ranking kept almost none: 0.1 for betweenness and 0 for degree, k-shell, PageRank and collective influence. Central users rarely leave, so support spent on them is wasted, even after the hub shock. Rankings built to find influential spreaders (Kitsak et al, 2010; Morone and Makse, 2015) answer a different question from whom to keep.

6 Shocks and interventions on Bluesky (RQ5)

The testbed starts at month 12 with the 2,106 agents active then. It runs under the selected independent-cascade mechanism and, for comparison, under the independent mechanism.

Table 10 Paired cascades and structure after shocks (30 seeds). Extra departures: departures among agents not removed by the shock, beyond the same agents' departures without the shock, also as a percentage of the shock size (105 agents). The difference with random removal is tested with a Wilcoxon signed-rank test over seeds (*** $p < 0.001$). LCC: largest component relative to its starting size; Recip.: reciprocity of the remaining network.

Mechanism	Shock (5%)	Extra dep.	vs. random	LCC	Recip.
Indep. cascade (selected)	Random	13.5 (13%)	–	0.88	0.30
	Hubs	51.0 (49%)	+37.5***	0.81	0.35
	Supportive	47.1 (45%)	+33.6***	0.83	0.23
Independent	Random	0	–	0.90	0.30
	Hubs	0	–	0.85	0.35
	Supportive	0	–	0.87	0.22

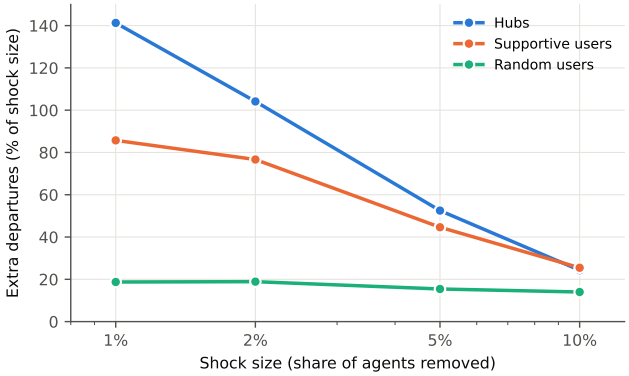

Fig. 4 Extra departures as a share of the shock, by shock size, under the independent cascade (15 seeds). Small losses of hubs are amplified most.

6.1 Shocks and cascades

Removing 5% of agents changed structure in the same way under both mechanisms (Table 10). Removing hubs did the most damage: the largest component shrank to 81–85% of its starting size, against 88–90% after random removal. Removing supportive users left the remaining network's reciprocity at 0.22–0.23, against 0.30 after random removal. Ties became more one-way.

Behavioral cascades depended on the mechanism. Under the independent mechanism, the paired cascade was exactly zero in every run, as designed. Under the independent cascade, hub removal caused 51 extra departures, 49% of the shock. Removing supportive users caused 47 (45%) and random removal 14 (13%). Hub and supportive-user shocks exceeded random ones by 37.5 and 33.6 departures ($p < 0.001$).

Amplification fell as shocks grew (Figure 4). Losing 1% of agents as hubs caused extra departures of 141% of the shock; at 10%, it caused 24%. Random shocks stayed between 14% and 19%. The first hubs to go are the most connected, so each one exposes many neighbors; later removals reach users whose neighbors have already been

Table 11 Intervention effects relative to the hub shock (S3), paired by seed (30 seeds). Kept: users kept over 12 months; Gain: retention gain in percentage points; p: Wilcoxon signed-rank; LCC gain: gain in the largest component's relative size, in points. Cost: ties added plus supported agent-months. IC: independent cascade.

Mech.	Intervention	Kept	Gain	p	LCC gain	Cost
IC	S5 Reconnection	−0.2	−0.01	0.059	5.0	504
(selected)	S6 Exposure-targeted	1.7	0.08	<0.001	0.0	504
	S7 Adaptive	0.2	0.01	0.098	2.8	1,007
	S8 Risk-targeted	**4.9**	**0.23**	<0.001	0.0	504
	S9 Oracle	6.3	0.30	<0.001	0.0	504
Indep.	S5 Reconnection	0.0	0.00	1.0	4.3	504
	S6 Exposure-targeted	2.0	0.10	<0.001	0.0	504
	S7 Adaptive	0.4	0.02	0.011	2.7	1,001
	S8 Risk-targeted	**4.9**	**0.23**	<0.001	0.0	504
	S9 Oracle	6.7	0.32	<0.001	0.0	504

exposed. Amplification also grew with social strength: 26%, 53% and 81% of a hub shock at 0.5, 1 and 1.5 times the fitted cascade probability. We saw no tipping point.

6.2 Interventions

With a budget of 2% of agents per month, every intervention had a small effect on a community of 2,106 (Table 11). The ranking was clear, though.

Reconnection repairs structure but keeps no one

Adding one tie per at-risk agent restored much of the connectivity the hub shock destroyed (5.0 points). Yet it kept no users under either mechanism. Under the independent cascade it was slightly harmful (−0.2 users, $p = 0.06$): every new tie is one more contact whose departure can pass on. A larger budget made this worse (−0.03 points at 5%). Structural repair and retention are different goals.

Whom to support matters more than how

Support for users with many departed contacts (S6) kept 1.7 users, because most of them were at low risk anyway. The same support for users with the highest M5 risk (S8) kept 4.9, 2.8 times as many under the independent cascade and 2.5 times under the independent mechanism. Risk targeting achieved 78% of the oracle's gain. The gap widened with the budget: at 5%, risk targeting gained 0.53 points against 0.19 for exposure targeting. Doubling the support effect raised its gain from 0.23 to 0.38 points. Section 5.4 showed that targeting by network centrality does far worse than both.

7 Discussion

7.1 What the framework shows

On Bluesky, departures cluster among contacts beyond what timing, activity and degree explain. A standard cascade model describes the clustering better than any

of the alternatives we tried, including SILCOM's own mechanisms. Under that model, losing the best-connected users sets off the most extra departures, and the first losses are amplified most. These results are consistent with a social process of withdrawal. They do not prove it. Homophily can produce the same clustering, as our benchmark and theory (Shalizi and Thomas, 2011) both show, and the best mechanism changed with the validation months.

Two findings hold whatever the mechanism. Departure is predictable from a user's own activity, as well as by a graph neural baseline. And support aimed at predicted risk keeps several times more users than support aimed at exposure, at random or at central users.

7.2 Lessons for building such systems

Several lessons apply beyond this study. *Compare against published methods inside the same pipeline.* Our own mechanisms lost to a classic cascade model from 2003; adding it changed the shock results more than any parameter sweep. *Separate detection from naming.* A permutation test detected social withdrawal reliably, but no method named the mechanism reliably, and no test ruled out homophily. *Measure cascades with paired runs.* Without common random numbers, random shocks appear to prevent departures. *Check the labels.* Defining withdrawal by inactivity at the window's end inflates late departure rates by 1.3 to 1.6 times even when the true rate is constant. *Validate targeting rules, not just predictors.* Centrality rankings that are standard for spreading problems failed completely for retention.

8 Situating the work in computational social science

We have framed this work for the social network analysis and mining community. The problem of withdrawal cascades, however, sits where several disciplines meet. Making those links explicit shows what the framework borrows and what it can give back.

Computational social science

Computational social science studies human behavior through large-scale digital traces and computational models. Community decline is one of its classic questions, studied through traces of who posts, replies and leaves (Kairam et al, 2012; Danescu-Niculescu-Mizil et al, 2013; Ribeiro, 2014). SILCOM adds two things. It turns observation into a testable model: the same traces calibrate a simulation whose mechanisms are scored on held-out months. And it moves from description to intervention: a community's own history becomes a testbed for responses before anyone acts. That reframes a familiar question—why do communities decline?—into a harder one: what would change the outcome, and for whom?

Network science

Withdrawal cascades are a network problem with a twist. Classic robustness work removes nodes and measures what is left (Tanizawa et al, 2005; Trajanovski et al, 2013). Classic diffusion work spreads a behavior along ties (Kempe et al, 2003). Here

the two meet: removal is itself the behavior that spreads. Our results add three network-science observations. Hub loss is amplified most, and small hub losses most of all. Adding ties restores connectivity without keeping people. And centrality rankings designed to find influential spreaders (Kitsak et al, 2010; Morone and Makse, 2015) are the wrong targets for retention. Decentralized platforms make the relevant graphs—replies, reposts, mentions and follows—publicly observable over time, which is rare.

Digital sociology and platform governance

Digital sociology asks how platform architecture shapes social life. Bluesky's architecture changes what leaving means. A user's records persist in a portable repository; communities can form around custom feeds and shared moderation (Kleppmann et al, 2024; Sokoto et al, 2026). Who keeps a community together, and who should act when it frays, are therefore questions for community organizers and moderators as much as for the platform. Our intervention results speak to that division of labor: the most useful signal for whom to reach is a member's own fading activity, which organizers can see, not their structural position.

Taken together, these links position withdrawal cascades as a shared object of study: a measurement problem for computational social science, a dynamical problem for network science and a governance question for digital sociology. Progress will need all three lenses at once.

9 Practical implications

The findings speak to three audiences.

Community organizers and moderators

Bluesky's open repositories let anyone run this kind of analysis on a community from public data alone. Organizers can watch for the departure of well-connected members, whose loss our simulations amplify most. They can direct scarce attention, such as welcomes, replies or invitations, to members whose own activity is fading, rather than to the most central members or to those whose friends have left. They should not expect new ties alone to stop a decline: reconnection repaired the network's structure but kept no one.

Platform and protocol designers

The AT Protocol already exposes what SILCOM needs: timestamped records, follows and interactions in public repositories. Designers of decentralized platforms can keep this openness while supporting community health, for example through community-level dashboards built on such signals, or opt-in tools that let organizers reach members at risk of drifting away. Because risk is best predicted from a user's own activity, such tools need little network data and can respect privacy.

Researchers

SILCOM is released with its benchmark and comparisons so that others can apply it to new communities. Two practices matter most. Report detection and mechanism separately, because the first is reliable and the second is not. And check end-of-window labeling bias whenever withdrawal is defined by inactivity.

Each of these is a hypothesis from a calibrated simulation. A field test would be needed before acting at scale.

10 Research agenda

Five directions follow for the social network analysis and mining community.

Identification

No test we compared separates contagion from homophily. Designs that exploit exogenous variation—platform outages, moderation events, policy changes or the departure of accounts for reasons unrelated to their contacts—could isolate social effects (O'Malley et al, 2014; Aral et al, 2009).

Generalization across communities

The same pipeline can be run on many Bluesky communities. Comparing them would show which findings are general and which belong to one community, and whether community structure predicts vulnerability once enough communities are observed.

Cross-platform comparison

Porting the collector to other decentralized platforms, such as Mastodon, would let researchers ask whether federated and protocol-based architectures differ in how communities decline. Mastodon's per-server data access requires per-server handling (Gao et al, 2025).

Validating retrospective reconstruction

Bluesky's live event stream records posts, deletions and account changes as they happen. Recording it prospectively for a sample and comparing it with a later retrospective collection would measure what reconstruction misses.

From simulation to the field

Risk-targeted support is the framework's strongest recommendation and also its least tested. Field experiments with community organizers, with consent and clear opt-outs, could test whether it keeps members in practice.

11 Limitations

Four limitations qualify our claims.

One community, one sample

The sample came from snowball sampling of one community from 12 seeds. Snowball samples favor well-connected users, and other communities on Bluesky may behave differently. Detecting clustering took about 2,500 eligible users; a subsample of 1,350 accounts had too little power.

Identification

Clustering beyond chance is not proof of contagion. Our benchmark shows that homophily can pass the permutation and shuffle tests. The selected mechanism also changed with the validation split, and two models had their best values at the edge of the grid.

Retrospective data

Retrospective collection sees only records that still exist. Deleted posts and deleted or deactivated accounts are invisible, so some departures may be missed or mistimed. The end-of-window labeling bias inflates late departures, and no fitted mechanism foresees the late rise.

Simulation

Simulated interventions are hypotheses. Their effects depend on the assumed support effect and on the mechanism; we report both mechanisms and sweep the main parameters to show how much.

12 Conclusion

SILCOM turns one collection of public Bluesky data into a rebuilt history, a test for clustered departures, a calibrated simulation and an intervention testbed. Each step is checked against ground truth and against published methods. Applied to 3,121 accounts, it found that departures cluster in the network and that a standard cascade model describes them best. Losing hubs is amplified most. Risk prediction from own activity matches a graph neural baseline, and support aimed at predicted risk keeps several times more users than support aimed at exposure, at random or at central users.

The take-home message is simple. Whom a community reaches matters more than how it reaches them, and the people worth reaching are those whose activity is fading, not those at the center. The framework also has clear limits. No method it contains, or that we compared, can separate contagion from homophily or reliably name the mechanism. Decentralized platforms make these questions answerable from public data for the first time. We release the framework, benchmark and comparisons so that others can apply them to new communities and push past these limits.

Supplementary information. Code, the synthetic benchmark, configuration files, the pseudonymized interaction data and all reported results are available for review at https://osf.io/6kdfj/overview?view_only=f85a5ff165d94b43a029e1f72a161c4a. A public version with a DOI will be released on publication.

Declarations

- **Funding.** This research received no specific funding.
- **Competing interests.** The author declares no competing interests.
- **Ethics approval.** Not applicable; this study uses only publicly available data from Bluesky.
- **Data availability.** The pseudonymized interaction data and all reported results are available at the OSF link above.
- **Code availability.** The SILCOM code, the synthetic benchmark and the configuration files for every result are available at the OSF link above.
- **Author contributions.** M. Al-Kfairy conceived the study, developed the framework, collected and analyzed the data, and wrote and revised the manuscript.
- **Use of generative AI.** During the preparation of this work, the author used generative AI tools (Claude, Anthropic) for proofreading, restructuring the text and some coding tasks. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Appendix A Parameters

Appendix B Robustness sweeps

Appendix C Additional results

References

Anagnostopoulos A, Kumar R, Mahdian M (2008) Influence and correlation in social networks. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, KDD '08, pp 7–15, https://doi.org/10.1145/1401890.1401897

Aral S, Muchnik L, Sundararajan A (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proceedings of the National Academy of Sciences 106(51):21544–21549. https://doi.org/10.1073/pnas.0908800106

Arroyo-Machado W, Robinson-Garcia N, Torres-Salinas D (2025) Are there stars in Bluesky? A comparative exploratory analysis of altmetric mentions between X and Bluesky. Journal of Informetrics 19(3). https://doi.org/10.1016/j.joi.2025.101700

Backstrom L, Huttenlocher D, Kleinberg J, et al (2006) Group formation in large social networks: Membership, growth, and evolution. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, KDD '06, pp 44–54

Balduf L, Sokoto S, Ascigil O, et al (2024) Looking AT the Blue Skies of Bluesky. In: Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC, pp

Table A1 Main parameters, all set in one configuration block of the released code.

Parameter	Value	Role
Observation window	18 months	Retrospective reconstruction
Calibration / validation	months 0–11 / 12–15	Held-out validation
Inactivity threshold τ	60 days (30, 90)	Withdrawal definition
Minimum actions	5	Eligibility
Linked-account rule	$\geq$ 20 shared timestamps	Cleaning
Network window (simulation)	4 months	Starting network
Permutations	200 (100 for the subsample)	Exposure test
Shuffles / bootstrap resamples	200 / 100	Shuffle and edge-reversal tests
Clustering weight λ	0.05	Calibration loss
Tie tolerance	1%	Mechanism selection
Calibration / validation runs	10 / 50 per setting	Simulation averaging
Horizon / seeds	12 months / 30	Scenario experiments
Shock size / budget	5% / 2% per month	S2–S9
Risk threshold / support effect / new ties	0.3 / 0.5 / 1	S5–S7
Prediction horizon	3 months	M5
Benchmark	6 mechanisms $\times$ 5 replicates $\times$ 1,500 users	System evaluation

76–91, https://doi.org/10.1145/3646547.3688407

Bektas A, Piana V, Schumann R (2021) A meso-level empirical validation approach for agent-based computational economic models drawing on micro-data: a use case with a mobility mode-choice model. SN Business and Economics 1(6). https://doi.org/10.1007/s43546-021-00083-4

Benlian A, Hess T (2011) The signaling role of IT features in influencing trust and participation in online communities. International Journal of Electronic Commerce 15(4):7–56. https://doi.org/10.2753/JEC1086-4415150401

Benlian A, Hoehne E, Hess T (2010) The contribution of it features to increase trust and participation in online communities: An empirical analysis. In: 18th European Conference on Information Systems, ECIS 2010

Bessems KMHH, Simovska V, Willems MDD, et al (2022) Factors influencing sustainability of online platforms for professionals: A mixed-method study in OECD countries. Health Promotion International 37(1). https://doi.org/10.1093/heapro/

Table B2 One-at-a-time sweeps under the independent cascade (15 seeds). Hub cascade: extra departures after the hub shock as a percentage of the shock. Gains: retention gain over S3 in percentage points. Default values in bold.

Parameter	Value	Hub cascade (%)	S5 gain	S6 gain	S8 gain
Social strength	0.5×	25.8	−0.010	0.092	0.231
	1×	52.5	−0.013	0.086	0.234
	1.5×	80.8	−0.016	0.095	0.234
Budget	1%	52.5	−0.006	0.044	0.127
	2%	52.5	−0.013	0.086	0.234
	5%	52.5	−0.028	0.190	0.532
Support effect u	0.25	52.5	–	0.035	0.136
	0.5	52.5	–	0.086	0.234
	1	52.5	–	0.146	0.380
Risk threshold	0.2 / **0.3** / 0.5	52.5	−0.013	0.086	–

Shock size	Random (%)	Hubs (%)	Supportive users (%)
1%	18.7	141.3	85.7
2%	18.9	104.1	76.7
5%	15.4	52.5	44.6
10%	14.0	24.3	25.5

Table C3 Network summary with and without quote posts (eligible users, whole window).

Measure	All interactions	Without quotes
Edges	65,105	59,066
Mean partners	25.8	23.4
Reciprocity	0.32	0.33
Mean clustering	0.32	0.31
Largest component share	0.94	0.93
Modularity	0.54	0.55
Top 1% share of interaction	0.38	0.39
Strength Gini	0.83	0.83

daab063

Beygelzimer A, Grinstein G, Linsker R, et al (2005) Improving network robustness by edge modification. Physica A: Statistical Mechanics and its Applications 357(3-4):593–612. https://doi.org/10.1016/j.physa.2005.03.040

Boero R, Squazzoni F (2005) Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. JASSS 8(4)

Table C4 Runtime (seconds) by module on synthetic data, and peak memory including the loaded case-study data.

Users	Events	Reconstruction	Analytics	Exposure model	One simulation run	Memory (GB)
1,000	86,622	0.1	1.0	0.2	0.1	2.0
2,500	209,860	0.3	5.4	0.2	0.3	2.0
5,000	424,460	0.5	8.4	1.5	0.9	2.0
10,000	855,101	1.6	21.2	0.5	1.2	2.2
20,000	1,703,351	3.0	53.5	0.9	3.7	2.4
50,000	4,254,620	5.9	169.7	2.3	9.7	2.9

Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30(1–7):107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Centola D, Macy M (2007) Complex contagions and the weakness of long ties. American Journal of Sociology 113(3):702–734. https://doi.org/10.1086/521848

Chen Y, Nevskaya Y (2026) Consumer Interactions and Peer Effects in Socially Connected Digital Products. Journal of Marketing Research https://doi.org/10.1177/00222437261463518

Chiu CM, Huang HY, Cheng HL, et al (2015) Understanding online community citizenship behaviors through social support and social identity. International Journal of Information Management 35(4):504–519. https://doi.org/10.1016/j.ijinfomgt.2015.04.009

Cox DR (1972) Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological) 34(2):187–220

Danescu-Niculescu-Mizil C, West R, Jurafsky D, et al (2013) No country for old members: User lifecycle and linguistic change in online communities. In: Proceedings of the 22nd International Conference on World Wide Web. ACM, WWW '13, pp 307–318

Dasgupta K, Singh R, Viswanathan B, et al (2008) Social ties and their relevance to churn in mobile telecom networks. In: Advances in Database Technology - EDBT 2008 - 11th International Conference on Extending Database Technology, Proceedings, pp 668–677, https://doi.org/10.1145/1353343.1353424

Diakidis G, Karna D, Fasarakis-Hilliard D, et al (2015) Predicting the evolution of communities in social networks. In: ACM International Conference Proceeding Series, https://doi.org/10.1145/2797115.2797119

Dileo M, Zignani M (2025a) Network-wide shocking events through the lens of node representation shift. In: CEUR Workshop Proceedings

Dileo M, Zignani M (2025b) User migration in blockchain-based online social networks through the lens of temporal node representation shift. Machine Learning 114(12).

https://doi.org/10.1007/s10994-025-06905-y

Ding J, Gao D, Chen X (2015) Alone in the game: Dynamic spread of churn behavior in a large social network a longitudinal study in MMORPG. International Journal of Smart Home 9(3):35–44. https://doi.org/10.14257/ijsh.2015.9.3.04

Dror G, Pelleg D, Rokhlenko O, et al (2012) Churn prediction in new users of Yahoo! Answers. In: Proceedings of the 21st International Conference on World Wide Web, Companion Volume. ACM, WWW '12 Companion, pp 829–834

Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35–41. https://doi.org/10.2307/3033543

Friedkin NE (1982) Information flow through strong and weak ties in intraorganizational social networks. Social Networks 3(4):273–285. https://doi.org/10.1016/0378-8733(82)90003-X

Galdeman A, Zignani M, Gaito S (2023) User Migration Across Web3 Online Social Networks: Behaviors and Influence of Hubs. In: IEEE International Conference on Communications, pp 5595–5601, https://doi.org/10.1109/ICC45041.2023.10278763

Gao M, Du H, Wen W, et al (2025) FediData: A Comprehensive Multi-Modal Fediverse Dataset from Mastodon. In: CIKM 2025 - Proceedings of the 34th ACM International Conference on Information and Knowledge Management, pp 6372–6376, https://doi.org/10.1145/3746252.3761634

Garcia D, Mavrodiev P, Schweitzer F (2013) Social resilience in online communities: The autopsy of Friendster. In: COSN 2013 - Proceedings of the 2013 Conference on Online Social Networks, pp 39–49, https://doi.org/10.1145/2512938.2512946

Guo M, Liu Y (2026) Online game player churn prediction based on multiplex social influence through network embedding. PLOS One 21(8). https://doi.org/10.1371/journal.pone.0356200

Hamilton WL, Zhang J, Danescu-Niculescu-Mizil C, et al (2017) Loyalty in online communities. In: Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017, pp 540–543

Han YJ, Moon J, Woo J (2024) Prediction of Churning Game Users Based on Social Activity and Churn Graph Neural Networks. IEEE Access 12:101971–101984. https://doi.org/10.1109/ACCESS.2024.3429559

Hironaka S, Yoshida M, Shudo K (2024) Comparing User Activity on X and Mastodon. In: Proceedings - 2024 IEEE International Conference on Big Data, BigData 2024, pp 2967–2972, https://doi.org/10.1109/BigData62323.2024.10825212

Huitema IR, Oskooei AR, Aktaş MS, et al (2025) Investigating Echo-Chambers in Decentralized Social Networks: A Mastodon Case Study. In: Studies in Computational Intelligence, pp 316–328, https://doi.org/10.1007/978-3-031-82431-9_26

Hwang S, Rollins S, Andrade Nunes T, et al (2026) Governing Together: Toward Infrastructure for Community-Run Social Media. In: Conference on Human Factors in Computing Systems - Proceedings, https://doi.org/10.1145/3772318.3790855

Ittefaq M (2025) Digital disconnective practice: Online platform migration and technology non-use in the age of emerging social media and polarized societies. Media, Culture and Society 47(4):815–823. https://doi.org/10.1177/01634437251326397

Jamal Z, Bucklin RE (2006) Improving the diagnosis and prediction of customer churn: A heterogeneous hazard modeling approach. Journal of Interactive Marketing 20(3-4):16–29. https://doi.org/10.1002/dir.20064

Jeong U, Nirmal A, Jha K, et al (2024) User Migration across Multiple Social Media Platforms. In: Proceedings of the 2024 SIAM International Conference on Data Mining, SDM 2024, pp 436–444, https://doi.org/10.1137/1.9781611978032.51

Joyce E, Kraut RE (2006) Predicting continued participation in newsgroups. Journal of Computer-Mediated Communication 11(3):723–747. https://doi.org/10.1111/j.1083-6101.2006.00033.x

Kairam SR, Wang DJ, Leskovec J (2012) The life and death of online groups: Predicting group growth and longevity. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining. ACM, WSDM '12, pp 673–682

Kang KW, Shin SK (2006) A model for the antecedents of member loyalty: The case of USENT newsgroup. In: Association for Information Systems - 12th Americas Conference On Information Systems, AMCIS 2006, pp 4500–4508

Kavak H, Padilla JJ, Lynch CJ, et al (2018) Big data, agents, and machine learning: Towards a data-driven agent-based modeling approach. In: Simulation Series, pp 125–136

Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, KDD '03, pp 137–146, https://doi.org/10.1145/956750.956769

Kimitei S, Agiro D, Ni S, et al (2025) Predictability & explainability of survival analysis in churn prediction. Journal of Marketing Analytics https://doi.org/10.1057/s41270-025-00450-2

Kitsak M, Gallos LK, Havlin S, et al (2010) Identification of influential spreaders in complex networks. Nature Physics 6(11):888–893. https://doi.org/10.1038/

nphys1746

Kleppmann M, Frazee P, Gold J, et al (2024) Bluesky and the AT Protocol: Usable Decentralized Social Media. In: DIN 2024 - Proceedings of the ACM Conext-2024 Workshop on Decentralization of the Internet, Co-Located with: CoNEXT 2024, pp 1–7, https://doi.org/10.1145/3694809.3700740

La Cava L, Greco S, Tagarelli A (2022) Information consumption and boundary spanning in Decentralized Online Social Networks: The case of Mastodon users. Online Social Networks and Media 30. https://doi.org/10.1016/j.osnem.2022.100220

La Cava L, Aiello LM, Tagarelli A (2023) Drivers of social influence in the Twitter migration to Mastodon. Scientific Reports 13(1). https://doi.org/10.1038/s41598-023-48200-7

La Cava L, Mandaglio D, Tagarelli A (2024) Polarization in Decentralized Online Social Networks. In: Proceedings of the 16th ACM Web Science Conference, WebSci 2024, pp 48–52, https://doi.org/10.1145/3614419.3644013

Lee M, Woo J (2025) TempODEGraphNet: predicting user churn using dynamic social graphs and neural ODEs. PLOS ONE 20(6 June). https://doi.org/10.1371/journal.pone.0321560

Li CT, Lin YJ, Yeh MY (2015) The roles of network communities in social information diffusion. In: Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015, pp 391–400, https://doi.org/10.1109/BigData.2015.7363780

Lu Z, Ali KS, Zhang Y, et al (2026) CARE: An Integrated Framework for Contagion-Aware Churn Prediction and Network-Level Risk Quantification. IEEE Access 14:103048–103070. https://doi.org/10.1109/ACCESS.2026.3703967

Malgonde OS, Padmanabhan B, Mithas S (2025) How to Mitigate Misinformation Diffusion on Blockchain-Based Decentralized Social Network Platforms? Insights From an Agent-Based Simulation Model. Journal of Operations Management 71(7):1068–1084. https://doi.org/10.1002/joom.70007

Manzo G, Gabbriellini S, Roux V, et al (2018) Complex Contagions and the Diffusion of Innovations: Evidence from a Small-N Study. Journal of Archaeological Method and Theory 25(4):1109–1154. https://doi.org/10.1007/s10816-018-9393-z

Mastroeni L, Naldi M, Vellucci P (2020) Calibration of an agent-based model for opinion formation through a retweet social network. In: CEUR Workshop Proceedings, pp 161–173

Melder E, Lerner A, Devito MA (2025) "A Blocklist is a Boundary": Tensions between Community Protection and Mutual Aid on Federated Social Networks. Proceedings of the ACM on Human-Computer Interaction 9(2). https://doi.org/10.1145/

3710919

Morone F, Makse HA (2015) Influence maximization in complex networks through optimal percolation. Nature 524(7563):65–68. https://doi.org/10.1038/nature14604

Muralidharan R, Ahn YY, Truong BT (2026) Federating Governance: How Community Rules Scale with Mastodon Servers. Proceedings of the ACM on Human-Computer Interaction 10(6). https://doi.org/10.1145/3816935

Myneni S, Sridharan V, Cobb N, et al (2018) Content-sensitive characterization of peer interactions of highly engaged users in an online community for smoking cessation: mixed-methods approach for modeling user engagement in health promotion interventions. Journal of Participatory Medicine 10(3). https://doi.org/10.2196/jopm.9745

Ng YMM, Ray R (2025) The journalists' exodus: Navigating the transition from Twitter to Mastodon and other alternative platforms. New Media and Society https://doi.org/10.1177/14614448251321165

Nicholson MN, Keegan BC, Fiesler C (2023) Mastodon Rules: Characterizing Formal Rules on Popular Mastodon Instances. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, pp 86–90, https://doi.org/10.1145/3584931.3606970

Nogara G, Sahneh ES, DeVerna MR, et al (2026) A longitudinal analysis of misinformation, polarization and toxicity on Bluesky after its public launch. Online Social Networks and Media 51. https://doi.org/10.1016/j.osnem.2025.100342

Oentaryo RJ, Lim EP, Lo D, et al (2012) Collective churn prediction in social network. In: Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2012, pp 210–214, https://doi.org/10.1109/ASONAM.2012.44

O'Malley AJ, Elwert F, Rosenquist JN, et al (2014) Estimating peer effects in longitudinal dyadic data using instrumental variables. Biometrics 70(3):506–515. https://doi.org/10.1111/biom.12172

Pavlopoulou MEG, Tzortzis G, Vogiatzis D, et al (2017) Predicting the evolution of communities in social networks using structural and temporal features. In: Proceedings - 12th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2017, pp 40–45, https://doi.org/10.1109/SMAP.2017.8022665

Petrič G (2015) Communicatively Integrated Model of Online Community: A Conceptual Framework and Empirical Validation on a Case of a Health-Related Online Community. https://doi.org/10.1515/9783110473780-007

Phan TM, Yarosh S (2016) Sustaining reciprocity: Generating social capital within peer-support communities. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, pp 369–372, https://doi.org/10.1145/2818052.2869115

Platt D (2020) A comparison of economic agent-based model calibration methods. Journal of Economic Dynamics and Control 113. https://doi.org/10.1016/j.jedc.2020.103859

Poiitis M, Vakali A, Kourtellis N (2022) On the aggression diffusion modeling and minimization in Twitter. ACM Transactions on the Web

Priester C, Schmitt S, Peixoto TP (2014) Limits and trade-offs of topological network robustness. PLoS ONE 9(9). https://doi.org/10.1371/journal.pone.0108215

Quelle D, Bovet A (2025) Bluesky: Network topology, polarization, and algorithmic curation. PLoS ONE 20(2 February). https://doi.org/10.1371/journal.pone.0318034

Radivojevic K, Adams D, Laszlo G, et al (2025) User migration in the Twitter diaspora. EPJ Data Science 14(1). https://doi.org/10.1140/epjds/s13688-025-00552-y

Ribeiro B (2014) Modeling and predicting the growth and death of membership-based websites. In: Proceedings of the 23rd International Conference on World Wide Web. ACM, WWW '14, pp 653–664

Sahneh ES, Nogara G, DeVerna MR, et al (2025) The Dawn of Decentralized Social Media: An Exploration of Bluesky's Public Opening. In: Lecture Notes in Computer Science, pp 422–437, https://doi.org/10.1007/978-3-031-78541-2_26

Santos T, Burghardt K, Lerman K, et al (2020) Can badges foster a more welcoming culture on Q&A boards? In: Proceedings of the 14th International AAAI Conference on Web and Social Media, ICWSM 2020, pp 969–973

Schwitter N (2024) How offline meetings affect online activities: the case of Wikipedia. EPJ Data Science 13(1). https://doi.org/10.1140/epjds/s13688-024-00506-w

Seidman SB (1983) Network structure and minimum degree. Social Networks 5(3):269–287. https://doi.org/10.1016/0378-8733(83)90028-X

Shalizi CR, Thomas AC (2011) Homophily and contagion are generically confounded in observational social network studies. Sociological Methods & Research 40(2):211–239. https://doi.org/10.1177/0049124111404820

Sittar A, Cesnovar M, Gucek A, et al (2026) Constructing a Dataset to Support Agent-Based Modeling of Online Interactions: Users, Topics, and Interaction Networks. IEEE Access 14:52890–52910. https://doi.org/10.1109/ACCESS.2026.3679263

Sokoto S, Balduf L, Ascigil O, et al (2026) Open or Blocked Skies? Community Moderation Practices in Bluesky. In: WWW 2026 - Proceedings of the ACM Web Conference 2026, pp 4505–4516, https://doi.org/10.1145/3774904.3792106

Stanoevska-Slabeva K, Schmid BF (2001) A typology of online communities and community supporting platforms. Proceedings of the Hawaii International Conference on System Sciences p 162. https://doi.org/10.1109/HICSS.2001.927041

Struett T, Sinnreich A, Aufderheide P, et al (2024) Can This Platform Survive? Governance Challenges for the Fediverse. International Journal of Communication 18:5607–5628

Tanizawa T, Paul G, Cohen R, et al (2005) Optimization of network robustness to waves of targeted and random attacks. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 71(4). https://doi.org/10.1103/PhysRevE.71.047101

Tausczik YR, Dabbish LA, Kraut RE (2014) Building loyalty to online communities through bond and identity-based attachment to sub-groups. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW, pp 146–157, https://doi.org/10.1145/2531602.2531688

Tkacz N, Cámara-Menoyo C, Zhang F (2026) Eventful migration: Rethinking social media migration with help from Elon Musk's sink. New Media and Society 28(8):4072–4094. https://doi.org/10.1177/14614448251358846

Trajanovski S, Martín-Hernández J, Winterbach W, et al (2013) Robustness envelopes of networks. Journal of Complex Networks 1(1):44–62. https://doi.org/10.1093/comnet/cnt004

Tsugawa S, Niida S (2019) The impact of social network structure on the growth and survival of online communities. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019, pp 1112–1119, https://doi.org/10.1145/3341161.3343526

Tsugawa S, Niida S (2021) Analyzing the effects of social network structure on the growth and survival of online communities in reddit. IEICE Transactions on Communications E104B(7):760–769. https://doi.org/10.1587/transcom.2020CQP0006

Ugander J, Backstrom L, Marlow C, et al (2012) Structural diversity in social contagion. Proceedings of the National Academy of Sciences of the United States of America 109(16):5962–5966. https://doi.org/10.1073/pnas.1116502109

Walk S, Helic D, Geigl F, et al (2016) Activity dynamics in collaboration networks. ACM Transactions on the Web 10(2). https://doi.org/10.1145/2873060

Wan R, Levonian Z, Yarosh S (2020) How much is a "like" worth? engagement and retention in an online health community. In: Proceedings of the ACM Conference

on Computer Supported Cooperative Work, CSCW, pp 413–417, https://doi.org/10.1145/3406865.3418320

Wu F, Souza A, Zhang T, et al (2019) Simplifying graph convolutional networks. In: Proceedings of the 36th International Conference on Machine Learning, Proceedings of Machine Learning Research, vol 97. PMLR, pp 6861–6871

Wu S, Das Sarma A, Fabrikant A, et al (2013) Arrival and departure dynamics in social networks. In: WSDM 2013 - Proceedings of the 6th ACM International Conference on Web Search and Data Mining, pp 233–242, https://doi.org/10.1145/2433396.2433425

Yu T, Thieme A, Zhang L (2025) Identifying Roadblocks in Building Decentralized Apps. In: ANRW 2025 - Proceedings of the 2025 Applied Networking Research Workshop, pp 168–170, https://doi.org/10.1145/3744200.3744778

Zhang G, Zeng J, Zhao Z, et al (2022) A counterfactual modeling framework for churn prediction. In: WSDM 2022 - Proceedings of the 15th ACM International Conference on Web Search and Data Mining, pp 1424–1432, https://doi.org/10.1145/3488560.3498468

Zhang Z, Lu M (2026) When engagement meets departure: how social capital, social engagement and social support impact churn within online health communities. Aslib Journal of Information Management 78(4):877–899. https://doi.org/10.1108/AJIM-04-2024-0350

Zignani M, Gaito S, Rossi GP (2018) Follow the "mastodon": Structure and evolution of a decentralized online social network. In: 12th International AAAI Conference on Web and Social Media, ICWSM 2018, pp 541–550