Third Party Effect: Community Based Spreading in Complex Networks

Shubham Agarwal
shubham.agarwal@hpe.com

Hewlett Packard Enterprise

June 12, 2017

Hewlett Packard Enterprise

Introduction

- One defining aspect of network structure is the presence of clusters of nodes which are commonly understood as sets of nodes that contains more connections among themselves than to the rest of the network.
- The goal is to study the properties of spreading processes where homophily affects not only the network structure but also spreading probabilities.

Algorithm

Algorithm 1 Community Based SIR Algorithm

Require: Pre Infected nodes $I_0 = \{v\}$, Graph $\mathcal{G} \equiv (\mathcal{E}, \mathcal{V})$,
Require: Λ, Ω, ζ

Initialize I with all the pre-infected nodes.
Initialize Π for all the nodes of the network.
while $|I| > 0$ **do**
    if $\zeta = 0$ **then**
        $break$
    end if
    $\zeta = \zeta - 1$
    Extract node v from I
    for each $u \in \Gamma(v)$ **do**
        if $\Pi(u) = 1$ **then**
            $p \sim (0, 1)$
            if $p < \Lambda_{\sigma(v),\sigma(u)}$ **then**
                $\Pi(u) \leftarrow 0$
                Add node u to I
            end if
        end if
    end for
    $q \sim (0, 1)$
    if $q < \Omega(v)$ **then**
        $\Pi(v) \leftarrow 2$
    else
        Add node v to I
    end if
end while

Case Study: Synthetic Networks

The main concern of this paper is to study the interaction among different communities in the context of heterogeneous spreading processes. We focus our study on two major aspects:

- **Intra-community Spreading:** Describes spreading processes within communities, which are governed by the diagonal elements of the spreading matrix $\lambda_{c,c}$
- **Inter-community Spreading:** Describes spreading processes across distinct communities, which are governed by the off diagonal elements of the spreading matrix $\lambda_{c,c'}$

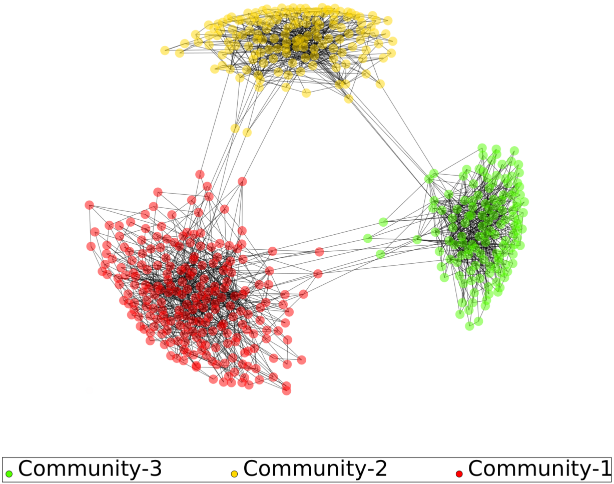
We have analyzed how the interplay between two communities which is essentially an inter-community process, is affected by intra-community interactions.

Case Study: Synthetic Networks

A hypothetical scenario has been considered in which interactions among three different communities **C0**, **C1**, **C2** have been examined:

- We have used a network of 500 nodes per community. Each community is generated as a complex network following a power law degree distribution $P(k) \sim k^{-3}$.
- The **C0** community is defined as the community in which a spreading process starts,
- **C1** is the target community where we quantify the impact of the infection in this community
- **C2** is the third party community whose inter-community spreading will come under scrutiny

Sample Graph



Impact Analysis

- In order to quantify the impact of the source community **C0** on the target community **C1**, we define the epidemic potential of infection of a community (*EPo*).
- *EPo* value has been calculated by performing 100 simulations with a given density of initially infected nodes in **C0** and then counting the number of simulations that cause at least half of the nodes in **C1** to be infected.
- We have studied the time difference between infection peaks to characterize the dynamical behavior of an epidemic under the heterogeneous framework. We define the average peak delay *APd* as $\langle | D_{\mathbf{C0}} - D_{\mathbf{C1}} | \rangle$ where D_c is the time at which the maximum amount of nodes are infected in community c.

Epidemic

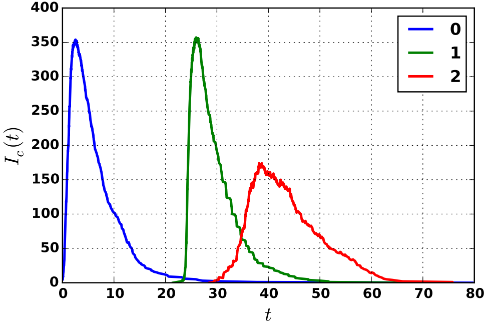
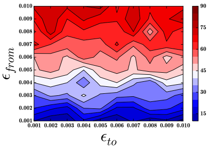
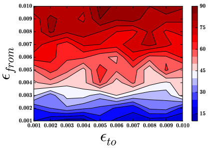


Figure: Example of an infection profile in real time with $\epsilon_{to} = 0.008$ and $\epsilon_{from} = 0.004$, $\lambda_{2,2} = 0.1$ showing the number of infected nodes for tree communities **C0**, **C1** and **C2** that form a synthetic network. See text for details.

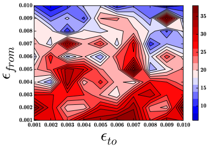
Phase Diagrams



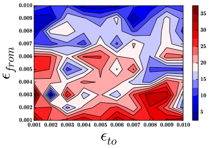
(a) Epo vs ϵ_{to} and ϵ_{from} for $\lambda_{2,2} = 0.1$. Lower intra communiy spreading rate for third party



(b) EPo vs ϵ_{to} and ϵ_{from} for $\lambda_{2,2} = 0.9$, Higher intra community rate for third party



(c) $\mid D_{\mathbf{C0}} - D_{\mathbf{C1}} \mid$ vs ϵ_{to} and ϵ_{from} for $\lambda_{2,2} = 0.1$, Lower intra communiy spreading rate for third party



(d) $\mid D_{\mathbf{C0}} - D_{\mathbf{C1}} \mid$ vs ϵ_{to} and ϵ_{from} for $\lambda_{2,2} = 0.8$, Higher intra community rate for third party

Real World Results

Name	# Communities	# Edges	# Nodes	*Epo*(1)	*Epo*(2)	$N(C_1)/N(C_0)$	$N(C_2)/N(C_0)$
CA-GrQc	421	14484	5242	2	10	0.87	0.36
CA-HepPh	416	118489	12008	98	98	0.92	0.32
CA-HepTh	546	25973	9877	14	14	0.95	0.93
Enron	1589	183831	36692	24	26	0.81	0.48
p2p-Gnutella04	25	39995	10876	6.67	13.33	0.96	0.78
p2p-Gnutella05	22	31840	8846	73.33	100.00	0.99	0.76
p2p-Gnutella06	20	31526	8717	26.67	60.00	0.83	0.77
p2p-Gnutella08	22	20778	6301	40.00	60.00	0.95	0.70
p2p-Gnutella09	30	26014	8114	20.00	73.33	0.89	0.78
p2p-Gnutella24	50	65370	26518	26.67	20.00	0.98	0.72
p2p-Gnutella25	58	54706	22687	6.67	6.67	0.86	0.83

Table: Results on real world networks from the SNAP data set. $N(C_i)/N(C_0)$ is the ratio of sizes of the target (third party) community to the source community. For source and target community we used an intra community spreading of $\lambda_{c,c} = 0.1$. (1) Stands for a third party intra community spreading of $\lambda_{2,2} = 0.2$ and (2) $\lambda_{2,2} = 0.8$.

Conclusions

We aimed at answering the question: Do the internal behaviors of third parties affect the interaction between two communities?

- Dynamic phase diagrams shows that third parties do effect the spreading between two communities.
- Enhancing effect through third parties can also be observed in networks with complex community structure.

Hewlett Packard
Enterprise

Fraunhofer
IAIS