Journal of Theoretical and Applied Information Technology
31st May 2022. Vol.100. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FITNESS PROPAGATION AND SELECTION OF SEED NODE FOR MESSAGE DIFFUSION IN SOCIAL NETWORKS

D.PUNITHA[1] , Dr.S.PRASANNA DEVI[2]

[1] Assistant Professor, SRM Institute of Science & Technology, Vadapalani Campus, Department of Computer Science, India

[2] HOD, SRM Institute of Science & Technology, Vadapalani Campus, Department of Computer Science, India

E-mail: [1] punithad@srmist.edu.in, [2] hod.cse.vdp@srmist.edu.in

ABSTRACT

In the social network, viral marketing, influence maximization is attained out by identifying influenced node in a network and attracting them by allocating allowances to propagate the message to their neighbors. The issue is this viral marketing is whether the message content propagated to users increases the utility of the users. Due to the practical importance of this problem to attain a win-win strategy, the problem has been analyzed in different works, and methodology has been proposed on how to attain Fitness in identifying seed node for message diffusion in social networks.

Keywords: *Viral Marketing, Utility Maximization, Message Propagation.*

1. INTRODUCTION

In current pandemic situation throughout world due to COVID-19, E-commerce retail is in high growth. Retailers follow different marketing strategy to attract the consumers. Viral marketing is the strategy followed in social network. The highly influenced social network user is identified and they are incentivized for forwarding the information to their neighbors

The interconnected structure of social networks enables social interaction between groups of agents. This interaction in social networks on a massive scale plays a significant role in sharing views and interests [21]. This interactive process plays a vital role in viral marketing. In viral marketing, the commercial production company does marketing by following different strategies to attract users for purchasing the product. The best way to do viral marketing is through influence maximization, where the highly influential set of users are identified and attract them with free samples. Each node in the influenced node-set will influence their neighbor nodes either by liking the product or by Word of mouth. These newly influenced users will, in turn, influence their neighbors by spreading the message about the product. This diffusion process of influencing neighbors will be continued and leads to significant improvement in the earned revenue of promoters [15]. During the influencing process, the number of free sample products depends on the budget allocated, and it will be limited in a wide-scale network [13]. Hence, to make an efficient process, the highly influential users are alone considered for distributing free samples based on the allocated budget. Based on this concept, the problem of selecting a highly influential set of users from the network is known as the *Social Influence Maximization Problem* [5].

Contribution

In existing works, the viral marketing of products involves multiple agents both in the merchant and consumer side. The agents can either be cooperative or competitive to accomplish a task. Based on game theory, a competitive agent environment attain win or lose strategy, whereas, in a cooperative environment, a win-win strategy is attained.

2. BACKGROUND

2.1 Influence Maximization and viral marketing

The fast-spreading of the message in the social network is due to multiple network links, which influence their neighbors through the diffusion of the message. This phenomenon of propagating the information to a subset of neighbors in the network, which may, in turn, transfer the message to their neighbors, and so on, is used by several stakeholders to promote their goals [7].

The growing trend in social media is to collect a huge volume of online data. The data is not just about the link data, which is linked with whom, but also about interaction data, which is interacting with whom [20]. Social network's ability of information diffusion rapidly in the network increases the business value and which in turn made the Influence Maximization (IM) an eye-catching task for product marketing. Viral marketing in social network adopts this IM methodology to identify influenced nodes. For viral marketing, the business marketing process selects a seed set of customers and activates each customer using free products, hoping that through Word of mouth effect, the product adoption would propagate in the network [8].

2.2 Multi-Agent Reinforcement Learning

Software agents are computer programs that help to bring in an adaptive process over different tasks of buying and selling products over the social network [1].
Bot: A bot is an autonomous agent that acts autonomously in creating 'new' data as they execute. For example, social networking bots may weave connections of interaction with a social graph.

Action: An action of the agent is a potentially state-changing operation performed on the platform.

Event: An event is a platform state change as a result of action taken by an agent that is visible to some set of users.

Observation: It is context observation made by agents. It purely functional which is decomposed into action that brings platform state changing

Read-only bot: a read-only bot is one that can only observe the state and cannot perform any actions.

Writer bot: a writer bot performs an action and, hence makes state change of platform on which it acts.

Social Reinforcement Learning is a sub-class of Multi-Agent Reinforcement Learning (MARL) where a large number of Reinforcement Learning (RL) based autonomous agents are related and interact with each other [16]. The objective of an agent in a real-world social network is to learn a policy and to capture inter-agent dependencies that map the network state of N users to the actions of N users. Thus, for an agent to learn the policy for a single user, it's required to consider the actions of all N users, which results in at least N2 parameters to learn the policy of N user with N actions per policy [3].

2.3 Q- Learning Algorithm

In Reinforcement Learning, an autonomous agent interact with the dynamic environment by communicating actions and rewards of its new state. Based on this interaction, agents learn a good control policy depending on the earned cumulative numerical rewards. Quality of learning was improvised using Multi –Armed Bandit algorithm to pick an action randomly [18]. This random picking is not an optimized method. Q-Learning is model-free online RL that learns the environment at runtime dynamically. Q-Learning algorithm is efficient in finding the best action rewards and selection policies. The expected future based reward for various state-action is represented as Q-Value. The agent interacting with the environment learns Q value over time S. Q-table is the implementation of the Action-Value function [19].

Self-adapting RL-based agents also termed as controllers which can plan for future based cumulative rewards. The learning of the RL-based controller happens by associating various system states and actions with rewards [10]. The controller exploits the learning by exploring the environment by performing utility maximization and reconfiguration of resources.

2.4 Utility Maximization

The utility measurement form user's view is done by calculating the variation between the user's valuation for the item and the price spent by the user on that item. Two sets of items are maintained by every node – desire set and adoption set. The desired set of any node is the set of items that the node has been informed about the seeding of its neighbors [16]. The adoption set of any node is the subset of adopted nodes from its desired set of nodes. The adopted nodes are a subset that maximizes the user's utility [25]. Hierarchical teaching is incorporated for learning agents in cooperative multi-agent environment, through which utility maximization has been attained [2]. But still some agents are not truthful and hold their capacity as private information which affects utility of users [12].

3. UTILITY MAXIMIZATION PROBLEM - VARIANTS

In this section, we have described a survey of relevant topics like influence maximization, viral marketing, utility maximization, multi-agent reinforcement learning, and different methodologies.

Social network-based utility maximization by exploring the preferences was propounded in [23], where each node download data by accessing to an expensive link. The other way Nodes can also acquire the universe at low cost is by using interlinked nodes through which can exchange copies of data segments among themselves. During this data over inter-node links, nodes in the group will not be truthful by hiding its capacity for its self-utility maximization. Such nodes are termed as 'non-reciprocating nodes', and these nodes are prohibited, such by proposing the “Give-and-Take” (GT) criterion. In this proposed GT concept, exchange among the nodes is allowed only when each participating node is truthful at least by sharing one segment to the other node, which is not available with the node. Based on GT criteria, utility maximization of each node depends on the availability of segments with the node .when establishing a link between two nodes. Both participating node's acceptance is required to establish a link. Unpaired nodes in the network use an expensive link to download data segments. Based on *the Stable Roommates Problem, Linear* complexity decentralized algorithms are used by nodes for choosing the best strategy depending on available information. The performance of a social group consisting of unpaired nodes is benchmarked using *the price of the Choice* method. The give and Take method is used to restrict *non-reciprocating behavior* in social groups to study the capacity of agents. Decentralized algorithms are used for deciding strategies of nodes with preferential exploration. In the absence of a facilitator between the nodes in social network *Decentralized Randomized algorithm* is used and is asymptotically optimal

Limitation:

Nodes are not truthful since there is a possibility to falsify information about the segment sets they have. All the proposed algorithms in this work, except decentralized randomized algorithms, assume that the information provided by each node has truthfulness. In such cases of no truthfulness of nodes, the performance of the nodes can degrade.

Maximizing Viral Advertising in Social Networks based on budget allocation is propound in [22].In influence maximization, the underlying assumption is that each user is incentivized based on the allocated budget without exceeding the allocated budget cost. However, the decision of the user for being initial adopters is not deterministic, but it's probabilistic. The utility model is used to characterize users’ satisfaction to understand how users make the decision for being initial adopters with the incentives within the allocated budget. In economics and game theory, the Utility function adopts the concept of non-linear concave function, which has been widely adopted for the measure of preferences over goods and services.

Based on two phases of Viral advertising, in the initial phase, the user may accept to be ab initial adopter based on the allocated budget. In the later phase, the advertising message will spread as a piece of information in social networks starting from the initial adopters. In this work, budget allocation in social networks is studied to maximize the spread of viral advertising. Utility functions are used to model the satisfaction of the user and the Independent Cascade (IC) model in spreading the advertisement. User satisfaction usually follows the rule of diminishing returns, where the decrease in the utility of user satisfaction is directly proportional to the increase of the budget, which satisfies the concave property of utility function. A greedy algorithm is used to select initial adopters. The budget allocation is done using a discrete greedy algorithm to compute the diffusion probability. This diffusion is done by adopting a Breadth-First search from the given node, which is the time-consuming process. In this algorithm, in each iteration, Budget elements are added uniformly. Utility gain in each node is “updated” instead of “computing” to avoid many traversals.

Limitation

The budget allocation in each iteration of the algorithm used is uniform. The marginal gain of each node is “updated” instead of “computing. Second, the spread of advertising in the social network is based on the diffusion models, and the content of the advertising is not considered, and the contribution of the users towards viral marketing based on the incentives they get is not considered.

In social networks, the competitive diffusion over the network is modeled based on Game-Theoretic Approach in [14]. The network model contains nodes that have independent tendencies with different levels of impact on other adjacent nodes. The goal of the players is to figure out the set of nodes and the content to ensure the sum of network tendencies shifts towards their side. While certain nodes are impacted with messages in the diffusion process, most of the nodes are not affected by the agents. As a result of this, the paper only labels a node as active if it can be impacted by a message. Models such as the linear threshold

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

model and the independent cascade model have been utilized to determine the Nash equilibrium.

Limitation:

The paper only focuses on the linear threshold model and independent cascade model, while a more exact solution might be attained by looking at the problem as an objective linear optimization. There also exist various different networks that are different from messaging networks.

Multi-armed mechanism-based quality-assuring is propounded in [11]. The paper focuses on finding an accurate cost-optimal way of selecting experts for specific tasks such that a minimum guaranteed amount of accuracy is maintained. The mechanism termed as Multi-armed bandit (MAB) has been discussed with the use of an Assured Accuracy Bandit (AAB) for quality assurance. The algorithms are further modified to realize a new algorithm, Constrained Confidence Bound for Strategic Setting (CCB-S). The CCB-S makes use of an ex-post monotone to learn about the experts and allows the accuracy level to be finalized with an optimal cost solution. The upper bound and the lower bound of the times the algorithms select a set that is sub-optimal is found to be bound by a constant factor.

Limitation:

While the target accuracy is guaranteed, Multi-armed bandit algorithms pick random subsets, and CCB-S holds high computational and optimization costs. Furthermore, the strategy employed by CCB-S makes sure that all experts are picked in the exploration steps ending up with very high costs depending on n value. The combinatorics involved in the problem prevents us from directly eliminating too many of the experts.

The multi-armed bandit auction mechanism for multi-unit procurement was propounded in [6]. The stochastic reward problem involving buyers in an auction trying to buy items from agents that are heterogeneous with two values involving cost and capacity is the basis of the model auction. The auction model bi-dimensional multi-armed bandit procurement discussed in this paper tries to increase the auctioneer expected utility to the maximum. This is done while keeping in mind the compatibility and individual rationality of the auctioneer. The quality of the agents is assumed to be known, and a mechanism for the auctioneer 2D-OPT is utilized to figure out the costs and capacities of the agents. A stochastic mechanism of BIC and IR, resulting in the learning mechanism 2D-UCB, is used to implement an allocation rule for learning unknown qualities of the agents. This helps the mechanism to procure one unit at a time to allocate appropriately with the qualities already learned.

Limitation

The model only focuses on a signal agent while a setting, including allocation over a set of agents, might be better. The learning algorithm could be completely characterized, as well.

In Competitive Influence Maximization, the Seed Selection based on budget allocated was propounded in [8]. The paper discusses action space as discrete and focuses on a two-phase budget allocation system by combining both seed and budget allocation. Influential nodes are convinced to act as a seed while the budget allocation is modified in a number of methods. 1) Consideration of influential nodes over other network nodes. 2) Discrete instead of continuous action space 3) IC, LT, and Triggering based diffusion models are used. TIM+ is used to identify the required nodes, and the values are estimated while the Double Oracle algorithms is used to figure out the Nash equilibrium.

Limitation:

The model is limited to only two players, and the consumer is ignored in this case.

Utility-Based Influence Maximization in viral marketing was propounded in [4]. The paper studies two different ways of sending a message through an initial set of nodes in a network. The propagation of the messages is studied while comparing their utility. The paper discusses the propagation vs. utility trade-off with respect to the message in the network. The goal is to pick two sets initially so that the sum of the utility value is maximum. The paper makes use of an efficient table-based algorithm for Bi-submodular functions to study the trade-off.

Limitation:

The utility function discussed in the paper is bi-submodular, which can be improved. The model can be improved to accommodate multiple users with varying utilities starting from the same message.

Summary Research Gap:

- To identify seed node for transmission of messages based on fitness of each node
- To ensure truthfulness in message diffusion
- To give reward only for truthful seed node

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- To achieve Nash equilibrium between seller and identified seed node

4. PROPOSED WORK:

An network is modeled as $G = (V, E)$. V is the set of nodes and E is the set of edges in the network. Directed Graph is used

Seed node selection:

For seed node selection strategy the accuracy of message propagation depends on the quality of selected seed node with truthfulness. In turn the selected seed node should be highly influenced in the core cluster identified
Seed Selection strategy

Community detection based on Jaccard Coefficient

$$J(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}$$

Cosine function to find the similarity between nodes

Cos(x, y) = x . y / ||x|| * ||y||
x . y = product (dot) of the vectors 'x' and 'y'.
||x|| and ||y|| = length of the two vectors 'x' and 'y'.
||x|| * ||y|| = cross product of the two vectors 'x' and 'y'

Table1 shows the Input Graph for Iteration 1 and Table 2 shows the Jaccard Coefficient to find Similarity between nodes. During iteration 1, cluster nodes are identified using cosine similarity as shown in Fig 1. The cosine similarity input is taken as input for Hierarchical Agglomerative Clustering, Fig 2 .The cluster The visualization of cluster is done using Gopher tool as shown in Fig 3.

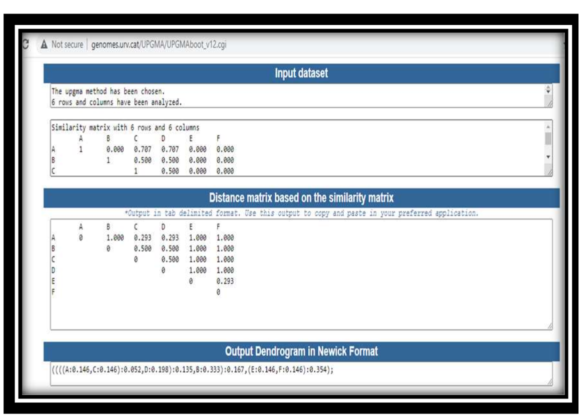

Fig:1: Hierarchical Agglomerative Clustering [HAC - Single Link]

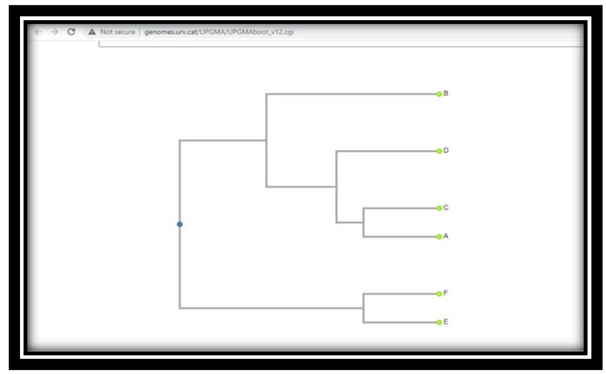

Fig 2: Hierarchical Agglomerative Clustering Output

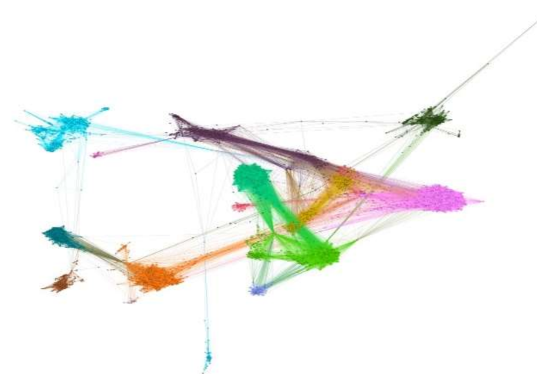

Fig 3: Cluster identification using Gephi tool

5. CONCLUSION:

In this work, works are reviewed based on the concepts of influence maximization, utility maximization, multi-agent-based reinforcement learning. The survey is carried out on utility maximization variants. Based on the literature review, various major challenges are identified and based on that proposed work to identify seed node based on fitness for message propagation. The proposed work reduces the time complexity of seed node identification. The limitations of proposed work does not include truthfulness of social network business, social platforms to identify seed node.

REFERENCES:

[1] John Ahlgren, Maria Eugenia Berezin, Kinga Bojarczuk, Elena Dulskyte, Inna Dvortsova, Johann George, Natalija Gucevska, Mark Harman, Ralf Lämmel, Erik Meijer, Silvia Sapora, Justin Spahr-Summers∗, (2020). "*WES: Agent-based User Interaction Simulation on Real Infrastructure*", ACM.

[2] Dong-Ki Kim, Miao Liu, Shayegan Omidshafiei, Sebastian Lopez-Cot, Matthew Riemer, Golnaz Habibi, Gerald Tesauro, Sami Mourad, Murray Campbell, Jonathan P. How, (2020). "*Learning Hierarchical Teaching Policies for Cooperative Agents"* arXiv:1903.03216v6 [cs. LG]. International Foundation for Autonomous Agents and Multiagent Systems.

[3] Mahak Goindani, Jennifer Neville (2020). "*Cluster-Based Social Reinforcement Learning*" arXiv:2003.00627v2 [cs. LG].

[4] Yael Sabato, Amos Azaria, Noam Hazon,(2020). "*Viral vs. Effective: Utility Based Influence Maximization*" International Foundation for Autonomous Agents and Multiagent Systems. ACM.

[5] Suman Banerjee, Mamata Jenamani, Dilip Kumar Pratihar, (2020). "*A survey on influence maximization in a social network*" Knowledge and Information Systems, Springer.

[6] Satyanath Bhat, Shweta Jain, Sujit Gujar, Y. Narahari, (2019). "*An optimal bidimensional multi-armed bandit auction for multi-unit procurement"* Annals of Mathematics and Artificial Intelligence. Springer.

[7] PRITHU BANERJEE, WEI CHEN, LAKS VS LAKSHMANAN, (2019). "*Maximizing Welfare in Social Networks under a Utility Driven Influence Diffusion Model"* ACM, arXiv:1807.02502v2 [cs. SI].

[8] Amirhossein Ansari, Masoud Dadgar, Ali Hamzeh, Jörg Schlötterer, Michael Granitzer, (2019). "*Competitive Influence Maximization: Integrating Budget Allocation and Seed Selection*" arXiv:1912.12283v1 [cs.SI].

[9] Orly Avner and Shie Mannor, (2019). "*Multi-User Communication Networks:A Coordinated Multi-Armed Bandit Approach"* IEEE/ACM TRANSACTIONS ON NETWORKING, Vol:27, 1063-6692.

[10] Dong-Ki Kim, Miao Liu, Shayegan Omidshafiei, Sebastian Lopez-Cot, Matthew Riemer, Golnaz Habibi, Gerald Tesauro, Sami Mourad, Murray Campbell, Jonathan P. How (2019). "*Heterogeneous Knowledge Transfer via Hierarchical Teaching in Cooperative Multiagent Reinforcement Learning*" Association for the Advancement of Artificial Intelligence.

[11] Shweta Jain, Sujit Gujar, Satyanath Bhat, Onno Zoeter, Y.Narahari, (2018). "*A quality assuring, cost optimal multi-armed bandit mechanism for expertsourcing*" Artificial Intelligence. Elsevier BV, 44-63.

[12] MICHAL ŠTOLBA, JAN TOŽIČKA, and ANTONÍN KOMENDA, (2018). "*Quantifying Privacy Leakage in Multi-Agent Planning*" ACM Transactions on Internet Technology, Vol. 18, No. 3, 1533-5399.

[13] Junqi Jin, Chengru Song, Han Li, Kun Gai, Jun Wang, Weinan Zhang (2018). "*Real-Time Bidding with Multi-Agent Reinforcement Learning in Display Advertising*" arXiv:1802.09756v2 [stat.ML].

[14] Shahla Jafari and Hamidreza Navidi (2018), "*A Game-Theoretic Approach for Modeling Competitive Diffusion over Social Networks*" Games, MDPI

[15] Suman Banerjeea, Mamata Jenamania, Dilip Kumar Pratihara (2018) "*A Survey on Influence Maximization in a Social Network*" arXiv:1808.05502v1 [cs. SI] Elsevier.

[16] Shayegan Omidshafiei, Jason Pazis, Christopher Amato, Jonathan P. How, John Vian (2017). "*Deep Decentralized Multi-task Multi-Agent Reinforcement Learning under Partial Observability*" PMLR 70.

[17] ANDREI MARINESCU, IVANA DUSPARIC, and SIOBHAN CLARKE', (2017). "*Prediction-Based Multi-Agent Reinforcement Learning in Inherently Non-Stationary Environments*" ACM Transactions on Autonomous and Adaptive Systems, Vol. 12, No. 2, 1556-4665.

[18] Shweta Jain, Sujit Gujar, Satyanath Bhat, Onno Zoeter, Y. Narahari, (2017). "*A quality assuring, cost optimal multi-armed bandit mechanism for expertsourcing*" Artificial Intelligence Elsevier BV, 44-63.

[19] HONGBIGN WANG, XIN CHEN, QIN WU, QI YU, XINGGUO HU, ZIBIN ZHENG, ATHMAN BOUGUETTAYA, (2017) "*Integrating Reinforcement Learning with Multi-Agent Techniques for Adaptive Service Composition*" ACM Transactions on Autonomous and Adaptive Systems, Vol. 12, No. 2, Article 8 1556-4665.

[20] Swapnil Dhamal, Rohith D. Vallam, and Y. Narahari, (2017). "*Modeling Spread of Preferences in Social Networks for Sampling-based Preference Aggregation*" TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, 2327-4697.

[21] ALEJANDRO TORREÑO, EVA ONAINDIA, ANTONÍN KOMENDA, MICHAL ŠTOLBA, (2017). "*Cooperative Multi-Agent Planning: A Survey*" ACM Computing Surveys, Vol. 50, No. 6, Article 84, 0360-0300.

[22] Bo-Lei Zhang, Wen-Zhong, Bin Tang, Sang-Lu Lu, Xiaoming Fu, (2016). "*Budget Allocation for Maximizing Viral Advertising in Social Networks*" J. Comput. Sci. & Technol. Springer, Vol: 31(4), 759-775.

[23] Saurabh Aggarwal, Joy Kuri, (2015). "*Strategies for utility maximization in social groups with preferential exploration*" Auton Agent Multi-Agent Syst, Springer.

[24] Qi-jin Ji, Zhe Yang, and Yan-qin Zhu, (2012). "*Utility-Maximizing Task Scheduling for Partially Observable Multiagent Systems*" AISC 125 Springer, 387-394.

Journal of Theoretical and Applied Information Technology
31st May 2022. Vol.100. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table1: Input Graph for Iteration 1

GRAPH	A	B	C	D	E	F	G	H	I
A	0	1	1	0	0	0	0	0	0
B	0	0	0	0	1	0	0	0	0
C	0	1	0	1	0	0	0	0	0
D	0	0	0	0	1	1	0	0	0
E	0	0	0	0	0	0	0	0	0
F	1	0	1	1	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0

Table 2:Jaccard Coefficient to find Similarity between nodes

JAC_COEFF	A	B	C	D	E	F
A	NA	0	0.5	0.5	0	0
B	NA	NA	0.333333	0.333333	0	0
C	NA	NA	NA	0.333333	0	0
D	NA	NA	NA	NA	0	0
E	NA	NA	NA	NA	NA	0.5
F	NA	NA	NA	NA	NA	NA