ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-1388

A Temporal-Causal Modelling Approach to Integrated Contagion and Network Change in Social Networks

Romy Blankendaal, Sarah Parinussa and **Jan Treur**[1]

Abstract. This paper introduces an integrated adaptive temporal-causal network model for dynamics in networks of social interactions addressing contagion between states, and changing connections within these social networks by two principles: the homophily principle and the more-becomes-more principle. The model has been evaluated in three different manners: by simulation experiments, by verification based on mathematical analysis, and by validation against an empirical data set.

1 INTRODUCTION

In today's world being successful and popular is mostly influenced by your social capabilities and how you interact with the people you know and work with. These social interactions are heavily investigated over the last few decades, in which analysis and prediction of the behaviour of humans in social situations plays a major role. The area of Social Networks has already a longer tradition, starting in the Social Sciences over 40 years ago. More recently, it has gradually developed in other disciplines as well; see, for example [4, 7, 23]. This development also involves computational methods to analyse and simulate networks both from the perspective of network structure and of dynamics.

Two main types of dynamics in relation to networks can be distinguished: dynamics *within* a given network structure (e.g., social contagion), and dynamics *of* a network (evolving networks). In the former case the network stays the same, but *states* (nodes) in the network change their *level* over time. In the latter case the network *connections* change, for example, their *weights* may increase or decrease. An example of this, also used in Section 6, is a network of adolescents followed over the years with the individuals' opinions about alcohol drinking as states and their friendships as connections. These states and connections both change over time. In many cases the two types of dynamics are addressed computationally as separate phenomena. This paper will address both types of dynamics and their interactions in an integrated manner.

The computational modelling approach used is the Network-Oriented Modelling approach based on temporal-causal networks described in [21, 22]; see also the ECAI'16 tutorial on Network-Oriented Modelling. This approach is a generic, dynamic AI modelling approach based on networks of causal relations but it differs from most other causal approaches (e.g., [17]) in that it incorporates a continuous time dimension to model dynamics in an adaptive manner, both of the states and of the network itself. This temporal dimension enables causal reasoning and simulation for cyclic and adaptive causal networks, such as networks for connected mental or brain states, or for social interactions, or both. The modelling approach can incorporate ingredients that are sometimes used in specific types of (continuous time, recurrent) neural network models, and ingredients that are often used in probabilistic or possibilistic modelling. It is more generic than such methods in the sense that a much wider variety of modelling elements are provided, enabling the modelling of many types of dynamical systems, as described in [21, 22].

For the dynamics within a network an existing approach to social contagion will be adopted. For the dynamics of the network, two different principles are considered and integrated, namely the *homophily* principle and the *more becomes more* principle. The main objective of this paper is to explore how combining of these three models for social contagion and network evolution can be used to analyse and predict human behaviour in social situations. The *contagion* principle indicates that the more a person interacts with someone else, the more their opinions or beliefs or emotions or other states will converge; e.g., [5]. The *homophily* principle in a sense indicates the converse of this: the similarity of states such as opinions or beliefs or emotions of persons affects the strength of the connection between them; e.g., [6, 12, 13, 19]. The *more becomes more principle* describes the phenomenon that whenever someone is popular (in the sense of the strength of connections), he or she will become more popular over time because it is believed that this person is worth relating to; e.g., [2, 15].

In this paper, in Section 2 the Network-Oriented Modelling approach based on temporal-causal modelling networks used is briefly described. Section 3 introduces the integrated network model, and in Section 4 simulation experiments are discussed. Section 5 addresses verification of the model by mathematical analysis of equilibria. In Section 6 validation of the model is addressed based on empirical data from [8]. Finally, a discussion is provided in Section 7.

2 TEMPORAL-CAUSAL NETWORK MODELS

Causal modelling, causal reasoning and causal simulation have a long tradition in AI; e.g., [10, 11, 17]. One of the challenges has been that causal modelling involving cyclic graphs is difficult; therefore, many approaches limit themselves to Directed Acyclic Graphs (DAG's). The Network-Oriented Modelling approach based on temporal-causal networks described in [21, 22] can be viewed as part of this tradition. The computational model presented here has been designed using this network modelling approach. It is a widely usable generic dynamic AI modelling approach that

[1] Computer Science Department, VU University Amsterdam, The Netherlands, romy.blankendaal@gmail.com, sparinussa@gmail.com, j.treur@vu.nl

distinguishes itself by incorporating a dynamic and adaptive perspective, both on states and causal relations. This dynamic perspective takes the form of an added continuous time dimension, enabling modelling of cyclic and adaptive networks, and also of timing of causal effects. Due to this, causal reasoning and simulation is possible for adaptive networks that inherently contain cycles, such as adaptive networks for connected mental or brain states, or for social interaction. From the technical point of view, it has some ingredients in common with specific types of (continuous time, recurrent) neural network models, but is more generic in the sense that a much wider variety of modelling elements can be used, as will also be shown by the integrated model presented here. In [21, 22] a more detailed description of this Network-Oriented Modelling approach can be found.

According to the adopted modelling approach, a model is designed at a conceptual level, for example, in the form of a graphical conceptual representation or a conceptual matrix representation. A graphical conceptual representation displays nodes for *states* and arrows for *connections* indicating causal impacts from one state to another, and includes some additional information in the form of:

- for each connection from a state X to a state Y a *connection weight* $\omega_{X,Y}$ (for the strength of the impact of X on Y)
- for each state Y a *speed factor* η_Y (for the timing of the effect of the impact)
- for each state Y the type of *combination function* $\mathbf{c}_Y(..)$ used (to aggregate multiple impacts on a state)

To choose combination functions, a number of standard options is available, varying from linear functions or logistic functions, to product or max and min-based functions as often used in probabilistic and possibilistic approaches; e.g., [21]. The conceptual representation of a model can be transformed in a systematic or even automated manner into a numerical representation of the model as follows [21]; here the variable t indicates a time point; it varies over the real numbers.

From conceptual to numerical representation

- at each time point t each state Y in the model has a real number value (usually in the interval [0, 1]), denoted by $Y(t)$
- at each time point t each state X connected to state Y has an impact on Y defined as $\mathbf{impact}_{X,Y}(t) = \omega_{X,Y} X(t)$ where $\omega_{X,Y}$ is the weight of the connection from X to Y
- The *aggregated impact* of multiple states X_i on Y at t is determined using a *combination function* $\mathbf{c}_Y(..)$:

  $\mathbf{aggimpact}_Y(t) = \mathbf{c}_Y(\mathbf{impact}_{X_1,Y}(t), \ldots, \mathbf{impact}_{X_k,Y}(t))$
  $= \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \ldots, \omega_{X_k,Y}X_k(t))$

  where X_i are the states with connections to state Y
- The effect of $\mathbf{aggimpact}_Y(t)$ on Y is exerted over time gradually, depending on speed factor η_Y:

  $Y(t+\Delta t) = Y(t) + \eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)] \Delta t$
  $\mathbf{d}Y(t)/\mathbf{d}t = \eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)]$
- This provides ase *difference* and *differential equation* for Y:

  $Y(t+\Delta t) = Y(t) + \eta_Y [\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \ldots, \omega_{X_k,Y}X_k(t)) - Y(t)] \Delta t$
  $\mathbf{d}Y(t)/\mathbf{d}t = \eta_Y [\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \ldots, \omega_{X_k,Y}X_k(t)) - Y(t)]$

These numerical representations can be used for mathematical and computational analysis and simulation.

In cases in which connection weights $\omega_{X,Y}$ are dynamic, they are also considered as states. This means that in graphical conceptual representations connection weights which usually are depicted as labels for arrows, can also be handled as nodes: also arrows can occur from and to them, as shown, for example in Fig. 1 and Fig. 2. For numerical representations dynamic connection weights also get a time argument: $\omega_{X,Y}(t)$. So the difference and differential equation for a state Y becomes:

$Y(t+\Delta t) = Y(t) + \eta_Y [\mathbf{c}_Y(\omega_{X_1,Y}(t)X_1(t), \ldots, \omega_{X_k,Y}(t)X_k(t)) - Y(t)] \Delta t$
$\mathbf{d}Y(t)/\mathbf{d}t = \eta_Y [\mathbf{c}_Y(\omega_{X_1,Y}(t)X_1(t), \ldots, \omega_{X_k,Y}(t)X_k(t)) - Y(t)]$

Moreover, as they are considered states themselves, to model their dynamics, the dynamic connection weights will also be described by a difference or differential equation, which also can be based on a combination function and speed factor as above, and even on weights of connections to or from these connection weights (the latter will be assumed to have value 1 here). This will be illustrated in detail in next section.

3 THE COMPUTATIONAL MODEL

Three different elements will be addressed by the computational model: the *contagion* principle, the *homophily* principle, and the *more becomes more* principle.

Contagion principle. This principle indicates that levels of states of connected nodes affect each other. A most basic form is that they are adjusted in a way that they become more equal, which can be considered a form of averaging; sometimes this is called absorption [5]. Also possible is that the level of one state amplifies the level of another state, so that spirals can occur; this is called amplification [5]. In the current paper an absorption model is used.

Homophily principle. This principle indicates that the more similar (the levels of) the states of two connected nodes are, the stronger their connection will become: 'birds of a feather flock together' [6, 12, 13].

More becomes more principle. This principle expresses that nodes that already have more and stronger connections get more and stronger additional connections than nodes with less or weaker connections (the rich become more rich and the poor remain poor). Analyses have been made showing that applying this principle usually leads to scale-free networks [2, 15]. When both the states and the connection weights are assumed dynamic, this leads to a circular causal relation state ⇄ connection. This may cause difficulties in explaining by which causes in the past observed phenomena in networks have developed. For example, when in a network it is found that similar state levels and strong connections occur together, due to such a circular causal relation it is difficult to tell which type of principle(s) was originally causing this situation; see for example: [1, 18, 20, 14].

3.1 Dynamics of States: Social Contagion

In this section, X_A denotes the level of state X for member A. This state X can be any type of state, either internal or externally observable (e.g. an internal state of feeling an emotion, or an expressed emotion state, or an (internal) intention, or an action performed, or a belief or opinion). The connection weights are denoted by $\omega_{A,B}$ for the connection from A to B. The following general model for contagion is used (see [21]):

$\mathbf{d}X_B/\mathbf{d}t = \eta_B [\, \mathbf{c}_B(\omega_{A_1,B}X_{A_1}, \ldots, \omega_{A_k,B}X_{A_k}) - X_B]$
$X_B(t+\Delta t) = X_B(t) + \eta_B [\, \mathbf{c}_B(\omega_{A_1,B}X_{A_1}(t), \ldots, \omega_{A_k,B}X_{A_k}(t)) - X_B(t)] \Delta t$

For the states X_B the model uses the scaled sum combination function (also see [5, 21]):

$\mathbf{c}_B(V_1, \ldots, V_k) = \mathbf{ssum}_\lambda(V_1, \ldots, V_k) = (V_1 + \ldots + V_k)/\lambda$

with $\lambda = \omega_{A_1,B} + \ldots + \omega_{A_k,B}$ the sum of the incoming weights for state X_B. This makes

$$c_B(\omega_{A_1,B} X_{A_1}(t), \ldots, \omega_{A_k,B} X_{A_k}(t)) = (\omega_{A_1,B} X_{A_1}(t) + \ldots + \omega_{A_k,B} X_{A_k}(t))/(\omega_{A_1,B} + \ldots + \omega_{A_k,B})$$

This is the weighted average of the state levels $X_{A_1}(t), \ldots, X_{A_k}(t)$ with weights proportional to $\omega_{A_1,B}, \ldots, \omega_{A_k,B}$, respectively. Using this combination function for the aggregated impact on state X_B, the differential and difference equation are

$$\mathbf{d}X_B/\mathbf{d}t = \eta_B [(\omega_{A_1,B} X_A) + \ldots + \omega_{A_k,B} X_{A_k})/(\omega_{A_1,B} + \ldots + \omega_{A_k,B}) - X_B]$$

$$X_B(t+\Delta t) = X_B(t) + \eta_B [(\omega_{A_1,B} X_{A_1}(t) + \ldots + \omega_{A_k,B} X_{A_k}(t))/(\omega_{A_1,B} + \ldots + \omega_{A_k,B}) - X_B(t)] \Delta t$$

3.2 Dynamics of Connections: Homophily

The network characteristics are specified by the connection strengths $\omega_{A,B}$. A first question to be answered is how these connection strengths are changing, and, in particular, the other states affecting them have to be identified. By the homophily principle the connection strengths $\omega_{A,B}$ are affected by the activation levels of the connected states of A and B. Such a dependency is depicted in fig. 1. Note that by adding these effects on the connection strengths cyclic relationships occur; for example see Figure 1.

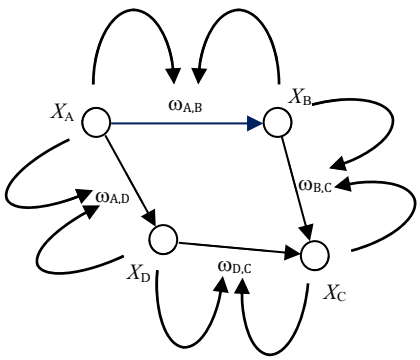

Figure 1. Conceptual representation: homophily principle

A next step is to determine how exactly the connection strengths are affected by the activation levels. This is needed to obtain a dynamic equation for $\omega_{A,B}$. For the current model the dynamic connection weights $\omega_{A,B}$ are assumed to change over time based on a principle similar to the one from [16]: the closer the activation levels of the states, the stronger the mutual connections between the members will become, and the higher the difference between the activation levels, the weaker they will become. In other words: activation levels close to each other imply a strong upward change in $\omega_{A,B}$, and activation levels far apart imply a downward change of $\omega_{A,B}$. This is how the homophily principle works: the more you are alike, the more you like (each other); the inspiration for the model below was obtained from [6, 12, 13, 19].

As an example of this principle in practical use, if you wonder whether there is a chance to become connected with somebody, you might consider whether you often like and agree on the same things. It can often be observed that persons that have close relationships or friendships are alike in some respects; e.g. they go to the same clubs, take the same drinks, have the same opinions, vote for the same or similar parties. Such observations might be considered support for the homophily principle: in the past they were attracted to each other due to being alike. However, also a different explanation is possible: they were often together and due to that they affected each other's states by social contagion, and therefore they became alike. So, the cyclic relation between X_B and $\omega_{A,B}$ as mentioned above leads to two possible causal explanations of a state of being alike and a state of being connected:

being connected $\rightarrow$ being alike
being alike $\rightarrow$ being connected

Such circular causal relations make it difficult to determine what came first. It may be a state just emerging from a cyclic process without a single cause. For more discussion on this issue, for example, see [1, 14, 18, 20].

The homophily principle may be formalised using a combination function $c_{A,B}(V_1, V_2, W)$ and speed factor $\eta_{A,B}$ according to the following general format (where connections to connection weights themselves are assumed to have weight 1):

$$\omega_{A,B}(t+\Delta t) = \omega_{A,B}(t) + \eta_{A,B} [c_{A,B}(X_A(t), X_B(t), \omega_{A,B}(t)) - \omega_{A,B}(t)] \Delta t$$

$$\mathbf{d}\omega_{A,B}/\mathbf{d}t = \eta_{A,B} [c_{A,B}(X_A, X_B, \omega_{A,B}) - \omega_{A,B}]$$

Here it is assumed that the values of $\omega_{A,B}$ stay within the interval [0, 1] and in particular the conditions

$$c_{A,B}(V_1, V_2, 0) \geq 0 \text{ and } c_{A,B}(V_1, V_2, 1) \leq 1$$

are fulfilled. The combination function $c_{A,B}(...)$ is assumed to depend on the one hand on W and on the other hand on the difference $|V_1 - V_2|$ (which is always between 0 and 1) in such a way that lower values of $|V_1 - V_2|$ relate to higher values of $c_{A,B}(V_1, V_2, W)$, and higher values of $|V_1 - V_2|$ relate to lower values of $c_{A,B}(V_1, V_2, W)$: the higher $|V_1 - V_2|$, the lower $c_{A,B}(V_1, V_2, W)$ and in particular:

$$|V_1 - V_2| = 1 \Rightarrow \omega_{A,B} \text{ decreasing } \Rightarrow \mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \leq 0$$
$$\Rightarrow c_{A,B}(V_1, V_2, W) \leq W$$
$$|V_1 - V_2| = 0 \Rightarrow \omega_{A,B} \text{ increasing } \Rightarrow \mathbf{d}\omega_{A,B}(t)/\mathbf{d}t \geq 0$$
$$\Rightarrow c_{A,B}(V_1, V_2, W) \geq W$$

Furthermore, it is assumed that $c_{A,B}(V_1, V_2, W)$ only depends on the difference $|V_1 - V_2|$ and not on the values of V_1 and V_2 themselves. Then as a simplification in notation the combination function $c_{A,B}(...)$ can be expressed as a function $h_{A,B}(D, W)$ of $D = |V_1 - V_2|$ and W: $c_{A,B}(V_1, V_2, W) = h_{A,B}(D, W)$. As discussed above, the function $h_{A,B}$ is assumed to be monotonically decreasing in D:

$$D_1 \leq D_2 \Rightarrow h_{A,B}(D_1, W) \geq h_{A,B}(D_2, W)$$
$$D = 1 \Rightarrow h_{A,B}(D, W) \leq W$$
$$D = 0 \Rightarrow h_{A,B}(D, W) \geq W$$

Moreover,

$$h_{A,B}(D, 0) \geq 0 \text{ and } h_{A,B}(D, 1) \leq 1$$

Somewhere between low values of $D = |V_1 - V_2|$ (with $h_{A,B}(D, W) \geq W$) and high values of D (with $h_{A,B}(D, W) \leq W$) a value for D is assumed for which $h_{A,B}(D, W) = W$. This is called the *homophily threshold value,* indicated by $\tau_{homophily}$ or by $\tau_{A,B}$; so

$$h_{A,B}(D, W) \geq W \text{ when } D \leq \tau_{A,B}$$
$$h_{A,B}(D, W) = W \text{ when } D = \tau_{A,B}$$
$$h_{A,B}(D, W) \leq W \text{ when } D \geq \tau_{A,B}$$

So, for this threshold value $\tau_{A,B}$ it holds:

- an upward change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| < \tau_{A,B}$
- no change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| = \tau_{A,B}$
- a downward change of connection weight $\omega_{A,B}$ occurs when $|V_1 - V_2| > \tau_{A,B}$

A simple example of a continuous function $h_{A,B}(D, W)$ satisfying the above conditions for a given value of W is obtained when the threshold value $\tau_{A,B}$ is assumed to be a fixed value and then use a simple decreasing linear function in D through the point with coordinates $(\tau_{A,B}, W)$ (i.e., through the point with $D = \tau_{A,B}$ and $h_{A,B}(D, W) = W$):

$$h_{A,B}(D, W) = W + \beta\,(\tau_{A,B} - D)$$

for some β, that still can be chosen. To fulfil the conditions

$$h_{A,B}(D, 0) \geq 0 \text{ and } h_{A,B}(D, 1) \leq 1$$

which prevent the weight value $\omega_{A,B}$ go outside the interval [0, 1], β can be chosen as a function $\beta(W)$ of W which can suppress the term $\tau_{A,B} - D$ when W comes closer to 0 or 1:

$$h_{A,B}(D, W) = W + \beta(W)\,(\tau_{A,B} - D)$$

When this function $\beta(W)$ is assumed to be always ≥ 0 and close to 0 when W is close to 0 or 1, then this can keep the value of $\omega_{A,B}$ within the interval [0, 1]. This can be satisfied by the function

$$\beta(W) = W\,(1\text{-}W)$$

which is 0 for $W = 0$ and for $W = 1$, and positive between these values with a maximum 0.25 for $W = 0.5$. This makes that ω is changing slowly in the neighbourhood of 0 or 1, thus achieving that ω does not cross these boundaries. Then the following example function fulfilling the above conditions is obtained:

$$h_{A,B}(D, W) = W + W\,(1\text{-}W)\,(\tau_{A,B} - D)$$

For the combination function $c_{A,B}(V_1, V_2, W)$ the above choice for $h_{A,B}(D, W)$ translates into:

$$c_{A,B}(V_1, V_2, W) = W + W\,(1\text{-}W)\,(\tau_{A,B} - |V_1 - V_2|)$$

Using this combination function, the dynamic relations for $\omega_{A,B}$ are:

$$\mathbf{d}\omega_{A,B}/\mathbf{d}t = \eta_{A,B}\,\omega_{A,B}(1 - \omega_{A,B})\,(\tau_{A,B} - |X_A - X_B|)$$

$$\omega_{A,B}(t+\Delta t) = \omega_{A,B}(t) + \eta_{A,B}\,\omega_{A,B}(t)(1 - \omega_{A,B}(t))\,(\tau_{A,B} - |X_A(t) - X_B(t)|)\,\Delta t$$

Note that in the example experiments discussed below one and the same homophily threshold value $\tau_{homophily}$ has been used for all connections.

3.3 Dynamics of Connections: More Becomes More

Another type of model for a dynamic connection from a member B to A takes into account to which extent other member's C connect to member A. The idea behind this is that somebody who is very popular seems worth connecting to. This is called the 'more becomes more' principle. For example, if B is followed by many others C on Twitter, then B seems to be interesting to follow for A as well. As the connections of others to B may change over time, this will imply that also A will have a dynamic connection to B, and in turn this connection will affect the connection of others to B over time as well. This can be modelled taking into account the weights $\omega_{C_i,B}$ for $i = 1,.., k$ of all connections from others C_i to B as follows. The inspiration for this model was obtained from [2, 15]. For a conceptual representation, see Fig. 2.

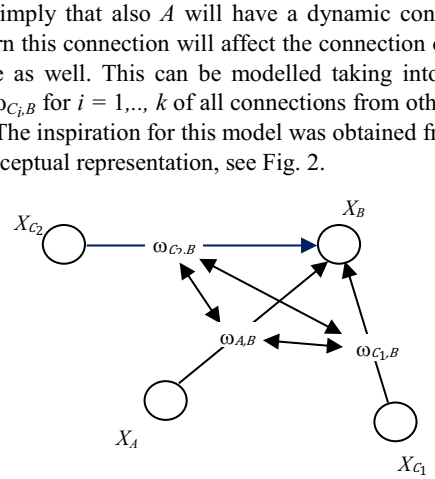

Figure 2. Conceptual representation: more becomes more principle

From this the following numerical representation is obtained:

$$\mathbf{d}\omega_{A,B}/\mathbf{d}t = \eta_{A,B}[c_{A,B}(\omega_{C_1,B} .., \omega_{C_k,B}) - \omega_{A,B}]$$

$$\omega_{A,B}(t+\Delta t) = \omega_{A,B}(t) + \eta_{A,B}[c_{A,B}(\omega_{C_1,B}(t), \ldots, \omega_{C_k,B}(t)) - \omega_{A,B}(t)]$$

Here $c_{A,B}(\ldots)$ is a combination function for the values $\omega_{C_1,B}, \ldots, \omega_{C_k,B}$, for example, a logistic sum function, or a scaled sum function. The latter is chosen here with as scale factor the number k of other members involved:

$$c_{A,B}(\omega_{C_1,B} .., \omega_{C_k,B}) = \mathbf{ssum}_k(\omega_{C_1,B}, .., \omega_{C_k,B}) = (\omega_{C_1,B} + .. + \omega_{C_k,B})/k$$

Note that in a network modelling the adaptation of connection weights the direction of this influence is not automatically the direction involving social contagion; this will depend on the application considered. For example, a network modelling a connection from A to B when *A is following B* on Twitter will not play a role in social contagion from A to B. For social contagion the opposite network plays a role where a connection from A to B occurs when *A is followed by B*, which is not initiated by A but by B: on Twitter and most other social media you cannot appoint your own followers. In other cases, it may be different. For example, if A wants to announce an event or new product, he or she can choose an occasion where many others will see the message, for example, posting it on a suitable forum; in such a case both the initiation and the social contagion are directed from A to the others. For the sake of simplicity, the latter is assumed here.

3.4 Integration of the Different Models

Within the integrated model the connection weights are affected by both the homophily principle and the more becomes more principle. These two effects have to be integrated. To achieve this a combination function is used that combines both the combination function for the homophily principle and for the more becomes more principle. Recall that the combination function for $\omega_{A,B}$ based on the homophily principle is:

$$\mathbf{c}_{\text{homo},\omega_{A,B}}(X_A, X_B, \omega_{A,B}) = \omega_{A,B} + \omega_{A,B}\,(1 - \omega_{A,B})*(\tau_{A,B} - |X_A - X_B|)$$

For the more becomes more principle the combination function for $\omega_{A,B}$ is:

$$\mathbf{c}_{\text{more},\omega_{A,B}}(\omega_{C_1,B}, .., \omega_{C_k,B}) = (\omega_{C_1,B} + .. + \omega_{C_k,B})/k$$

In order incorporate the influence of the separate models into the combined model, a parameter α between 0 and 1 is introduced and then for the integrated combination function $\mathbf{c}_{\omega_{A,B}}(\ldots)$ the weighted average is chosen of the two separate combination functions:

$$\mathbf{c}_{\omega_{A,B}}(X_A, X_B, \omega_{A,B}, \omega_{C_1,B}, .., \omega_{C_k,B}) = \alpha * \mathbf{c}_{\text{homo},\omega_{A,B}}(X_A, X_B, \omega_{A,B}) + (1\text{-}\alpha) * \mathbf{c}_{\text{more},\omega_{A,B}}(\omega_{C_1,B}, .., \omega_{C_k,B})$$

This is used as aggregated impact in the following difference for weight $\omega_{A,B}$:

$$\omega_{A,B}(t+\Delta t) = \omega_{A,B}(t) + \eta_{\omega_{A,B}}[\mathbf{c}_{\omega_{A,B}}(X_A(t), X_B(t), \omega_{A,B}(t), \omega_{C_1,B}(t), .., \omega_{C_k,B}(t)) - \omega_{A,B}(t)]$$

Parameter α is called the *homophily influence fraction*. When α is close to 1, the *homophily* model is dominating, and when α is close to 0, the *more becomes more* model dominates. When α is 0.5, both models have equal influence on the new weight value.

4 SIMULATION EXPERIMENTS

The model described above can be used to make predictions about emergent properties of dynamic social networks. To examine this, some questions concerning emerging properties were formulated and these were tested by simulation experiments. Due to the *contagion* principle it may be expected that the state levels of the nodes connected in a social network to become more alike and due to the *homophily* principle and the *more becomes more* principle their connections will change over time. Via the *more becomes more* principle it may be expected that the nodes that are connected to nodes that have the most and stronger connections will keep

those connections and they will become stronger. The homophily principle also changes the connections depending on the levels of the nodes. An example social network has been designed to see what the outcome is for these questions in a case study. The network contains 14 nodes and consists of two groups that each are strongly interconnected and have only a few (bridge) connections between members of different groups; see Fig. 3. These could, for example, be two groups of friends of different high schools that have a few connections between them due to four members that went to the same primary school. The assumed connection weights are shown in the connection matrix in Table 1.

So, will the integrated model change the levels of the states in each of the groups in such a way that they will converge to the same value for this group, depending on the popular nodes and the amount and strength of the connections? Or will the levels of all nodes converge to one and the same value for both groups? Besides that, due to the homophily principle, will the bridge connections between the two separate groups become weaker as their states are too different, or will they become stronger due to the more become more principle and due to the levels in the groups becoming more similar? Many types of emerging dynamic patterns may occur, and it is not so easy to predict them at forehand.

In the example about the high school friend groups, by the homophily principle the bridge connections could disappear when the persons in first group all play basketball and the people in the other friend group don't like basketball. However, when both groups like basketball, it is quite likely that the connections between the two groups will become stronger. And the more becomes more principle may still have a stronger effect to increase the bridge connections.

Table 1. Initial weights of the connections in the example network

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A		0.3	0.6	0.8										
B	0.5			0.6										
C				0.7										
D	0.6				0.8		0.3							
E				0.8			0.6	0.7	0.6			0.6		
F				0.3			0.9							
G				0.7		0.4		0.8						
H					0.7		0.9							
I										0.6	0.6	0.9		
J									0.7		0.8			
K									0.8	0.7			0.6	
L					0.4			0.5	0.5				0.6	0.7
M									0.4					0.6
N					0.8							0.7	0.6	

4.1 Analysis of the Example Social Network

Gephi version 0.8.2. [3] has been used to analyse the example social network, which is shown in Figure 3. It consists of 14 nodes or nodes and 39 edges. The average degree is 2.786 with a highest in-degree of 6, which means that there is one node who has 6 connections directed towards him/her; this is node D. As can be seen from Fig. 3, the most 'popular' nodes in the network are nodes D, E, I and L, these are the biggest and the size here indicates the number of connections for nodes. Node E has the highest betweenness centrality of 80.167, after which node D follows with 72.5, node L with 41 and node I with 38.167. The highest values for the between-ness centrality are most likely the popular members and have the most influence on the rest of the connecting nodes. The number of communities is 3 and the modularity is 0.369. The example network was designed as having 2 communities with a few nodes that also had bridge connections with the other community: nodes E, H, I, L, N.

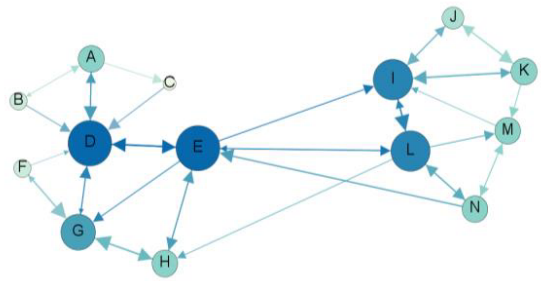

Figure 3. The example social network with nodes (circles) and edges (arrows with direction)

This group is also considered a community, probably since the modularity is 0.369 and the communities used here are small. For the two simulation experiments discussed here the parameter settings and initial values used are shown in Tables 1 to 3. Simulations were done in Matlab 2014v [4].

Table 2. Parameters used for the simulation experiments

Parameter	α	$\eta_{weights}$	η_{states}	$\tau_{homophily}$	time
Value	0.5	0.2	0.2	0.05	100

4.2 Simulation Results for Experiment 1

The first experiment addresses the case that the initial levels in the two groups do not differ much: the two groups have preferences about sports that match. The values of the levels of all nodes in the social network vary between 0.3 and 0.8. Because the levels are not so different, it may be expected that the bridge connections between the two groups (involving nodes E, H, I, L and N) will become stronger and the levels of all the nodes converge to the same value. Fig. 4 depicts the change of the levels of the nodes over time. The levels start at different initial values ranging from 0.3 to 0.8, as can be seen in Table 3. The graph shows that indeed all levels converge into one value, around 0.53.

Table 3. Initial values of the levels of the nodes

Node	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Experiment 1	0.4	0.6	0.3	0.7	0.5	0.6	0.4	0.8	0.5	0.3	0.6	0.6	0.7	0.4
Experiment 2	0.8	0.6	0.7	0.9	0.7	0.5	0.6	0.8	0.2	0.4	0.1	0.3	0.4	0.2

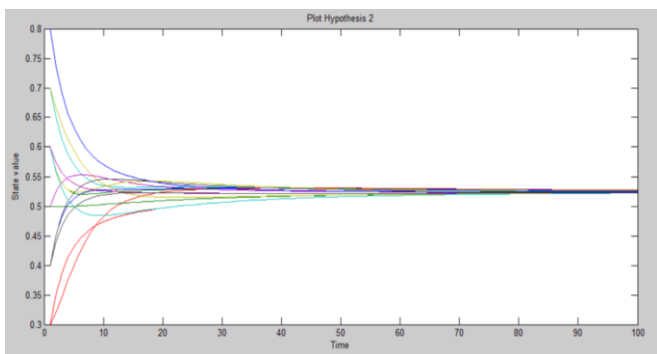

Figure 4. Experiment 1: changing levels of the states

In Fig. 5 the weights of the bridge connections between the two groups are shown which are expected to develop higher values. This is indeed the case; eventually they all converge to 1. This

happens because after time point 20 all levels differ less than 0.05, so both the homophily principle and the more becomes more principle have an increasing effect.

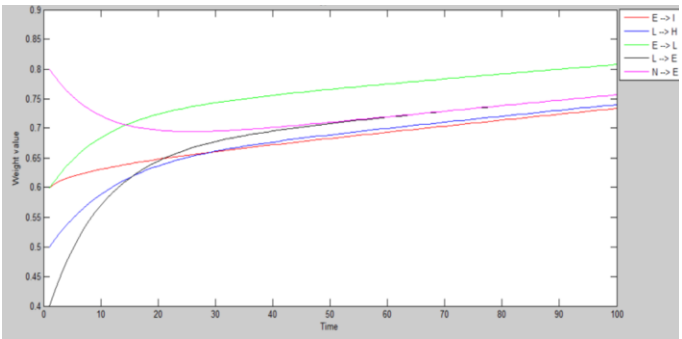

Figure 5. Experiment 1: changing bridge connection weights (between the nodes E, H, I, L and N)

4.3 Simulation Results for Experiment 2

In the second Experiment 1: changing bridge connection weights simulation experiment the two groups have their own opinion about sports. The first group, consisting of nodes A, B, C, D, E, F, G, H, considers basketball a really nice sport, therefore the initial levels of their states are high (values between 0.5 and 0.9). The second group, consisting of nodes I, J, K, L, M, N, have low initial levels as they do not like basketball (values between 0.1 and 0.4). One question was whether in some cases the bridge connections between nodes E, H, I and L will vanish over time due to the effect of homophily. Fig. 6 depicts the change of the levels of the nodes over time. As can be seen from the plot, the levels in both groups start of from the different initial values. Group 1 consists of the lines in the graph that start from 0.5 to 0.9 and group two consists of the lines starting below ranging from 0.1 to 0.4. The graph shows two things. First, there is an effect of increased clustering within each group (most in Group 2), but later the values of these clusters still converge to one value, around 0.53.

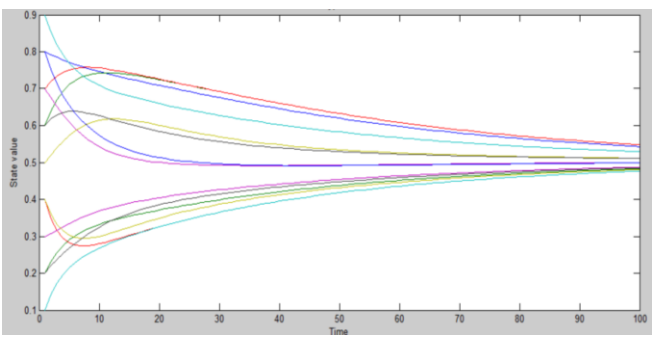

Figure 6. Experiment 2: changing levels of the states

So, in the longer term the bridge connections still make that one value is reached, which indicates that the bridge connections do not vanish. In Fig. 7 the weights of the bridge connections between the two groups are shown. A question was whether they would decrease to close to 0, due to the homophily principle. As can be seen from the plot, this is not the case. The connections between these nodes stay quite high and they are becoming higher over time. This can be explained by the more becomes more principle as follows. Note that Fig. 3 shows that the bridge connections all involve quite popular nodes. Therefore, the more becomes more principle makes these connections stronger instead of weaker, as would the homophily principle do. Apparently the more becomes more principle dominates in this case, and makes that effects of the homophily principle do not emerge. When after some time the differences between the levels become less than the homophily threshold $\tau_{homophily} = 0.05$, even the homophily principle starts to contribute to the increase of the connection weights.

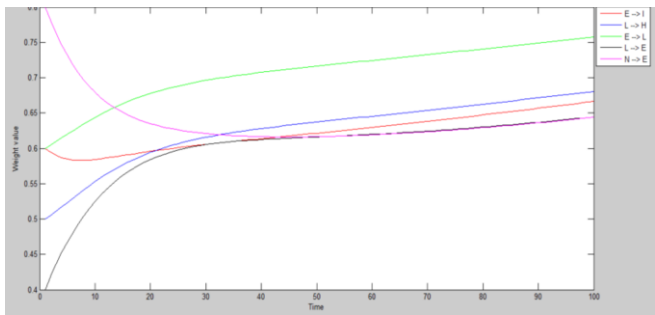

Figure 7. Experiment 2: changing bridge connection weights (between the nodes E, H, I, L and N)

5 VERIFICATION BY MATHEMATICAL ANALYSIS

This section presents some of the results of a mathematical analysis for the model. These results agree with the example simulations that have been performed, some of which are shown in Section 4. This provides verification of the model. First, a variable V indicating a state or connection weight is called *stationary* at t if $\mathbf{d}V(t)/\mathbf{d}t = 0$, it is *increasing* at t when $\mathbf{d}V(t)/\mathbf{d}t > 0$ and *decreasing* at t when $\mathbf{d}V(t)/\mathbf{d}t < 0$. For a connection weight $\omega_{A,B}$ being stationary means $\mathbf{d}\omega_{A,B}(t)/\mathbf{d}t = 0$; this is equivalent to:

$$\alpha\, \mathbf{c}_{\text{homo},\omega_{A,B}}(X_A(t), X_B(t), \omega_{A,B}(t)) + (1-\alpha)\, \mathbf{c}_{\text{more},\omega_{A,B}}(\omega_{C_1,B}(t), .., \omega_{C_k,B}(t)) = \omega_{A,B}(t)$$

which is modelled as

$$\alpha\, (\omega_{A,B}(t) + \omega_{A,B}(t)\,(1 - \omega_{A,B}(t))\,(\tau_{A,B} - |X_A(t) - X_B(t)|)) + (1-\alpha)\, (\omega_{C_1,B}(t) + .. + \omega_{C_k,B}(t))/k = \omega_{A,B}(t)$$

Similarly for states X_A. An *equilibrium* is when all states and connections are stationary at t. By rewriting the above formulae the following criteria can be derived for $\omega_{A,B}$ being stationary, increasing, and decreasing at time t.

Proposition 1 (Criteria for stationary, increasing, decreasing)

Stationary at *t*:

$$\alpha\, \omega_{A,B}(t)\,(1 - \omega_{A,B}(t))\,(\tau_{A,B} - |X_A(t) - X_B|(t)) + (1-\alpha)\,[\,(\omega_{C_1,B}(t) + .. + \omega_{C_k,B}(t))/k - \omega_{A,B}(t)] = 0$$

Increasing at *t*:

$$\alpha\ \omega_{A,B}(t)\,(1 - \omega_{A,B}(t))\,(\tau_{A,B} - |X_A(t) - X_B(t)|) + (1-\alpha)\, *\,[\,(\omega_{C_1,B}(t) + .. + \omega_{C_k,B}(t))/k - \omega_{A,B}(t)\,] > 0$$

Decreasing at *t*:

$$\alpha\, \omega_{A,B}(t)\,(1 - \omega_{A,B}(t))\,(\tau_{A,B} - |X_A(t) - X_B(t)|) + (1-\alpha)\, *\,[\,(\omega_{C_1,B}(t) + .. + \omega_{C_k,B}(t))/k - \omega_{A,B}(t)\,] < 0$$

Next, an equilibrium analysis will be made for the cases when $0<\alpha<1$. Later, in Proposition 4 the special case $\alpha = 0$ (more becomes more) will be addressed and in Proposition 5 the case $\alpha = 1$ (homophily). An equilibrium state can be considered for a connection weight $\omega_{A,B}$ for the three cases $\omega_{A,B} = 0$, $\omega_{A,B} = 1$ or $0 < \omega_{A,B} < 1$.

Proposition 2 (Equilibrium) *Let values* $\underline{X}_A$ *and* $\underline{\omega}_{A,B}$ *for all A and B for an equilibrium state be given. Then for all A and B the following hold:*

(a) *If* $\underline{\omega}_{A,B} = 0$ *and* $\alpha<1$, *then* $\underline{\omega}_{C,B} = 0$ *for all C connected to B*

(b) *If* $\underline{\omega}_{A,B} = 1$ *and* $\alpha<1$, *then* $\underline{\omega}_{C,B} = 1$ *for all C connected to B*

(c) *If* $0 < \underline{\omega}_{A,B} < 1$, *and* $\alpha>0$ *then*

$$|\underline{X}_A - \underline{X}_B| = \tau_{A,B} + (1/\alpha - 1) * [\ (\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B}\ ]/[\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})]$$

Proof (a) From $\underline{\omega}_{A,B} = 0$ and $\alpha<1$ it follows that $(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k = 0$, and since $0 \leq \underline{\omega}_{C,B} \leq 1$ for all C this entails $\underline{\omega}_{C,B} = 0$ for all C

(b) Similar to (a)

(c) This follows from:

$$\alpha * \underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ (\tau_{A,B} - |\underline{X}_A - \underline{X}_B|) = -(1-\alpha) * [\ (\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B}\ ]$$

$$\tau_{A,B} - |\underline{X}_A - \underline{X}_B| = -(1-\alpha) * [\ (\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B}\ ]/[\alpha\ \underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})]$$

$$|\underline{X}_A - \underline{X}_B| = \tau_{A,B} + (1/\alpha - 1) * [\ (\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B}\ ]/[\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})]$$ ■

More specifically, for an equilibrium in which all values X_A are the same the following is found.

Proposition 3 (Equilibrium with equal state values) *Assume* $0<\alpha<1$. *Let for an equilibrium state values* $\underline{X}_A$ *and* $\underline{\omega}_{A,B}$ *for all A and B be given, such that* $\underline{X}_A = \underline{X}_B$ *for all A and B. Then for each B the equilibrium values* $\underline{\omega}_{C,B}$ *for all C either are all equal to* 0 *or are all equal to* 1:

$$\underline{\omega}_{C,B} = 0 \text{ for all } C \text{ or } \underline{\omega}_{C,B} = 1 \text{ for all } C$$

Moreover, all connection weights $\omega_{A,B} > 0$ *are attracted to the equilibrium value* $\underline{\omega}_{A,B} = 1$, *and not to the value* $\underline{\omega}_{A,B} = 0$.

Proof From $\underline{\mathbf{X}_A = \mathbf{X}_B}$ it follows that

$$\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ (\tau_{A,B} - |\underline{X}_A - \underline{X}_B|) = \underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ \tau_{A,B} \geq 0$$

Take the smallest $\underline{\omega}_{A,B}$ for B. Then

$$(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B} \geq 0$$

Therefore since their sum is 0, both $\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ \tau_{A,B}$ and $(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B}$ are 0:

$\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B}) = 0$, which is equivalent to $\underline{\omega}_{A,B} = 0$ or $\underline{\omega}_{A,B} = 1$

$(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k = \underline{\omega}_{A,B}$

By Proposition 2 if $\underline{\omega}_{A,B} = 0$, then all $\underline{\omega}_{C,B} = 0$, and if $\underline{\omega}_{A,B} = 1$, then all $\underline{\omega}_{C,B} = 1$. So, either $\underline{\omega}_{C,B} = 0$ for all C or $\underline{\omega}_{C,B} = 1$ for all C.

The case $\underline{\omega}_{C,B} = 0$ for all C is not attracting: if for one of the $\underline{\omega}_{C,B} > 0$ and $\underline{\omega}_{C,B} < 1$, then $(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k - \underline{\omega}_{A,B} \geq 0$ and $\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ \tau_{A,B} > 0$, so it is increasing.

Some special cases have been excluded in parts of the above analysis: $\alpha = 0$ or $\alpha = 1$. When $\alpha = 0$ the model describes only the more becomes more principle. Then the equilibrium equations are:

$$\mathbf{c}_{\text{more},\omega_{A,B}}(\underline{\omega}_{C_1,B}, .. , \underline{\omega}_{C_k,B}) = \underline{\omega}_{A,B}$$

$$(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k = \underline{\omega}_{A,B}$$

For this case the following can be found.

Proposition 4 (More becomes more: $\alpha = 0$). *Assume* $\alpha = 0$ *and let equilibrium values* $\underline{\omega}_{C,B}$ *be given. Then for all C and D the equilibrium values* $\underline{\omega}_{C,B}$ *and* $\underline{\omega}_{D,B}$ *are equal:* $\underline{\omega}_{C,B} = \underline{\omega}_{D,B}$.

Proof Take the A such that $\underline{\omega}_{A,B}$ is the lowest from the $\underline{\omega}_{C,B}$. Then from

$$(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B})/k = \underline{\omega}_{A,B}$$

it follows

$$(\underline{\omega}_{C_1,B} + .. + \underline{\omega}_{C_k,B}) - k\,\underline{\omega}_{A,B} = 0$$

$$(\underline{\omega}_{C_1,B} - \underline{\omega}_{A,B}) + .. + (\underline{\omega}_{C_k,B} - \underline{\omega}_{A,B}) = 0$$

where each term $\underline{\omega}_{C_i,B} - \underline{\omega}_{A,B} \geq 0$. Therefore $\underline{\omega}_{C_i,B} = \underline{\omega}_{A,B}$ for all i. ■

Next, consider $\alpha = 1$, a model for only the homophily principle. Then the equilibrium equation becomes

$$\mathbf{c}_{\text{homo},\omega_{A,B}}(\underline{X}_A, \underline{X}_B, \underline{\omega}_{A,B}) = \underline{\omega}_{A,B}$$

$$\underline{\omega}_{A,B} + \underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ (\tau_{A,B} - |\underline{X}_A - \underline{X}_B|) = \underline{\omega}_{A,B}$$

$$\underline{\omega}_{A,B}(1 - \underline{\omega}_{A,B})\ (\tau_{A,B} - |\underline{X}_A - \underline{X}_B|) = 0$$

The solutions of this equation are:

$\underline{\omega}_{A,B} = 0$ or $\underline{\omega}_{A,B} = 1$ or $|\underline{X}_A - \underline{X}_B| = \tau_{A,B}$

Proposition 5 (Homophily: $\alpha = 1$). *Assume* $\alpha = 1$ *and let equilibrium values* $\underline{\omega}_{A,B}$ *be given. Then for all A and B it holds*

$$\underline{\omega}_{A,B} = 0 \quad \text{or} \quad \underline{\omega}_{A,B} = 1 \quad \text{or} \quad |\underline{X}_A - \underline{X}_B| = \tau_{A,B}$$

6 VALIDATION FOR REAL WORLD DATA

For validation, data have been used from [8] for a large network of adolescents at three time points in subsequent years at which measurements have been done about how the state or opinion about alcohol drinking is and how their friendships are. In total, there were 160 participants, but only 129 were present at all three measurements. For the network structure, the participants were asked to name a maximum of six friends. They were also asked about their behaviour, for example, about their alcohol drinking behaviour, whether they were using drugs and their smoking behaviour. The following data files from the collection were used:

Friendship 1: About the relationships between subjects at time point 1
Friendship 2: About the relationships between subjects at time point 2
Friendship 3: About the relationships between subjects at time point 3
Alcohol: The alcohol drinking behaviour of every subject at the 3 time points

The model uses continuous values between 0 and 1. Therefore, the data points from the dataset were transformed into values that are comparable with data that will be predicted by the model. The score 'best friend' has been mapped on value 0.9, 'just a friend' on 0.5 and 'no relation' on 0.1. The values are chosen in this way partly because the function used to determine the next value of a connection weight ω contains the factor $\omega(1-\omega)$. That means that when ω is exactly 0 or 1, no change can occur. By using the above values, this will not be a problem.

In order to compare the empirical data with the model data used in this project, simulations were performed for 20 times steps. At t = 0 the initial values were taken from the first time point of the empirical data, in this way giving the model the same starting position as the empirical data. Then the simulated values at time point 10 and at time point 20 were compared to the second and third time point of the empirical data. To compare the simulated data to the empirical data the parameters have been tuned to this specific network by Simulated Annealing (e.g., [9]); For the Simulated Annealing as an error function, the sum of squares of deviations between simulated model values and empirical values has been used, in order to get a minimal error. The parameters used

and their ranges are shown in Table 4. It was found out that the speed factors $\eta_{weights}$ and η_{states} should not be too high for proper functioning of the model.

As mentioned, the Glasgow data [8] was collected for three time points in subsequent years. There are cases that at the first time point there was no reported relation between two subjects, but at time point 3 there was. However, this is not taken into account by the model used here. The model does not create new connections. An alternative option could be to assume by default a very low weight initially, but that option has not been chosen. Therefore, a selection of the Glasgow data was made: relations between two subjects who did not yet have a relation during the initial time point, have been ignored. By applying the Simulated Annealing parameter estimation method to the empirical data and the model data, the parameter values shown in the rightmost column in Table 4 were identified.

Table 4. Parameters, their ranges and the identified values

Parameter	Notation	Interval	Value
Update speed factor for states	$\eta_{weights}$	[0, 0.55]	0.2811
Update speed factor for connections	η_{states}	[0, 0.25]	0.2065
Fraction of homophily influence	α	[0, 1]	0.2556
Threshold for homophily principle	$\tau_{homophily}$	[0, 1]	0.1945

The pattern for the error, defined as the average of the squares of the deviations between the empirical data and simulated data, is shown in Fig. 8. The lowest value is 0.0035; the square root of this value is 0.0592. Considered as a measure for deviations within the [0, 1] interval where the values of connections and states vary, this means an average deviation of 6%, which is not bad as an approximation; this provides a positive validation result.

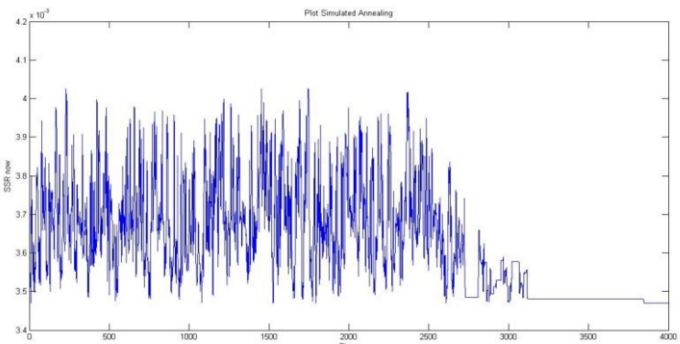

Figure 8. Simulated annealing: pattern of average of squares of deviations between empirical data and simulated data

In Fig. 9 for just two typical examples of connection weights it is shown how the dynamics of the model compares to the empirical values. Note that in Fig. 9 the (isolated) empirical values have been displayed with connecting lines between them; however, there is no linear relation, they are just measurements at different time points. Note that due to the discrete scores in the empirical data, real values can only be 0.1, 0.5 or 0.9. Here the real value of weight 1 stays 0.9 all the time, whereas the value of the model changes a bit, but deviates at most 0.02. The real value of weight 2 first stays 0.9 but then changes from 0.9 to 0.5 between the second and third time point. The value of the model also decreases, with deviations between 0.1 and 0.2.

Note that the identified value of the homophily influence fraction parameter α was around 0.25. This can be interpreted in the sense that according to the model in this network the evolution of relations is determined for about 25% by the homophily principle based on similarity in alcohol drinking, and for about 75% by the more becomes more principle.

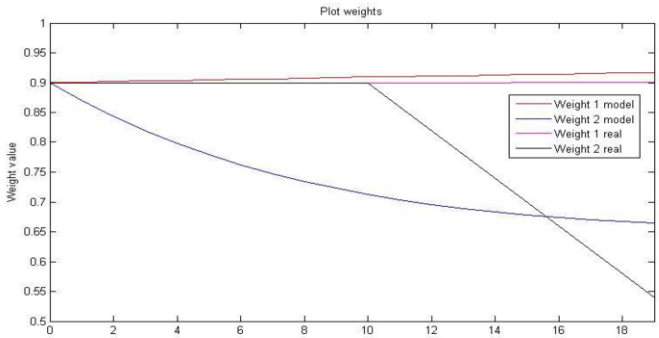

Figure 9. Two examples of the dynamics of connection weights in the model in comparison to the empirical values

7 DISCUSSION

The computational model for dynamics in social networks introduced in this paper was designed by integrating three models for the dynamics of states and relations of persons in a social network, addressing the *contagion* principle [5], the *homophily* principle [6, 12, 13, 19] and the *more-becomes-more* principle [2, 15], respectively. The model relates to and differs from existing work as follows. The first two models were adopted as variations on existing models from [5, 19], and the third model and the integration of the three models are new, as far as the authors know.

The integrated model was evaluated in three different manners: by various simulation experiments, by verification based on mathematical analysis, and by validation against an empirical data set. By all three methods a positive evaluation was found. The simulation results were as expected or at least were well explainable, the verification showed that the model provides the outcomes as predicted by the mathematical analysis, and the validation provided outcomes of the model with deviations from the empirical data within a 6% range of the [0, 1] interval. As part of the validation it was found that for the considered network of adolescents [8], according to the tuned model the dynamics of connections was determined by the homophily principle based on similarity in alcohol drinking for about 25%, and by the more becomes more principle for about 75%.

For future research variations of the model can be analysed. For example, for the homophily sub-model, as an alternative a quadratic variant as introduced in [19] can be used. Moreover, homophily with respect to multiple states can be explored, for example, not only similarities for drinking alcohol but also for smoking and sports behaviour. Also alternatives for the more becomes more and social contagion sub-models can be analysed, for example by choosing other combination functions from the collection of possible combination functions shown in [21].

Finally, validation can be performed for more empirical data on social networks and their dynamics over time. However, often data sets for social networks only provide data for one point in time and only about the connections. It is not easy to find data sets that include data about connections over time and in addition also data about states of the members of the network over time. In this sense the Glasgow data set [8] used here is very valuable and has a rather unique position.

REFERENCES

[1] S. Aral, , L. Muchnik, and A. Sundararajan, Distinguishing Influence Based Contagion From Homophily Driven Diffusion in Dynamic Networks. *Proceedings of the National Academy of Sciences* (USA), **106**, 1544-1549, (2009).

[2] A.L. Barabási, and R. Albert, Emergence of Scaling in Random Networks. *Science*, **286**, 509-512, (1999).

[3] M. Bastian, S. Heymann, and M. Jacomy, *Gephi: an open source software for exploring and manipulating networks*. https://gephi.org/, 2009.

[4] S. Boccalettia, V. Latorab, Y. Morenod, M. Chavez, and D.-U. Hwanga, Complex networks: Structure and dynamics. *Physics Reports,* **424**, 175 – 308, (2006).

[5] T. Bosse, R. Duell, Z.A. Memon, J. Treur, and C.N. van der Wal, Agent-Based Modelling of Emotion Contagion in Groups. *Cognitive Computation Journal*, **7**, 111-136, (2015).

[6] D. Byrne, The attraction hypothesis: Do similar attitudes affect anything? *Journal of Personality and Social Psychology*, **51**, 1167-1170, (1986).

[7] J. Giles, Computational Social Science: Making the Links. *Nature*, **488**, 448-450, (2012).

[8] Glasgow Empirical Data https://www.stats.ox.ac.uk/~snijders/siena/Glasgow_data.htm, 2016

[9] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi, Optimization by Simulated Annealing, *Science, New Series*, **220**, 671-680, (1983).

[10] B.J. Kuipers, Commonsense reasoning about causality: Deriving behavior from structure, *Artificial Intelligence,* **24,** 169-203, (1984).

[11] B.J. Kuipers and J.P. Kassirer, How to discover a knowledge representation for causal reasoning by studying an expert physician. In: *Proceedings Eighth International Joint Conference on Artificial Intelligence, IJCAI'83,* Karlsruhe. William Kaufman, Los Altos, CA, 1983.

[12] M. McPherson, , L. Smith-Lovin, and J.M. Cook, Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.,* **27**, 415–444, (2001).

[13] A. Mislove, B. Viswanath, K.P. Gummadi, and P. Druschel, You Are Who You Know: Inferring User Profiles in Online Social Networks. *Proc. WSDM'10*, 2010, New York City, New York, USA, pp. 251-260, 2010.

[14] M.P. Mundt, L. Mercken, and L.I. Zakletskaia, Peer selection and influence effects on adolescent alcohol use: a stochastic actor-based model. *BMC pediatrics*, **12**, 115, (2012).

[15] M.E.J. Newman, The structure and function of complex networks. *SIAM Review*, **45**, 167–256, (2003).

[16] H.V.D. Parunak, E. Downs, and A. Yinger, Socially-Constrained Exogenously-Driven Opinion Dynamics. *Proc. of the Fifth International IEEE Conference Self-Adaptive and Self-Organizing Systems (SASO'11)*, 2011.

[17] J. Pearl, *Causality*. Cambridge University Press, 2000.

[18] C.R. Shalizi, , and A.C. Thomas, Homophily and Contagion are Generically Confounded in Observational Social Network Studies. *Sociological Methods & Research,* **40**, 211–239, (2011).

[19] A. Sharpanskykh, and J. Treur, Modelling and Analysis of Social Contagion in Dynamic Networks. *Neurocomputing*, **146**, 140–150, (2014).

[20] C.E.G. Steglich, T.A.B. Snijders, and M. Pearson, Dynamic networks and behavior: separating selection from influence. *Sociol Methodol.,* **40**, 329–393, (2010).

[21] J. Treur, Dynamic Modelling Based on a Temporal/Causal Network Modelling Approach. *Biologically Inspired Cognitive Architectures Journal*, **16**, 131-168, (2016).

[22] J. Treur, *Network-Oriented Modelling: Addressing Complexity of Cognitive, Affective and Social Interactions*. Understanding Complex Systems Series, Springer Publishing, 2016.

[23] T.W. Valente, *Social Networks and Health: Models, Methods, and Applications*. New York, Oxford University Press, 2010.