Contagions across networks: colds and markets

Matthew J. Berryman[a], Neil F. Johnson[b] and Derek Abbott[a]

[a]Centre for Biomedical Engineering and School of Electrical and Electronic Engineering, The University of Adelaide, SA 5005, Australia.
[b]Physics Department, Clarendon Laboratory, Oxford University, Parks Road, Oxford OX1 3PU, United Kingdom

ABSTRACT

We explore a variety of network models describing transmission across a network. In particular we focus on transmission across composite networks, or "networks of networks", in which a finite number of networked objects are then themselves connected together into a network.

In a disease context we introduce two interrelated viruses to hosts on a network, to model the infection of hosts in a classroom situation, with high rates of infection within a classroom, and lower rates of infection between classrooms. The hosts can be either susceptible to infection, infected, or recovering from each virus. During the infection stage and recovery stage there is some level of cross-immunity to related viruses. We explore the effects of immunizing sections of the community on transmission through social networks.

In a stock market context we introduce memes, or virus-like ideas into a virtual agent-based model of a stock exchange. By varying the parameters of the individual traders and the way in which they are connected we are able to show emergent behaviour, including boom and bust cycles.

Keywords: small world networks, financial modeling, viruses

1. INTRODUCTION

1.1. SIRS model

An SIRS "household" model describes the spread of an infectious agent through a network with two types of connections between infectable agents:

1. local connections within a household that have a high probability of infection transmission and

2. connections between households that have a low probability of infection transmission.[1]

The three states of an SIR model are:

1. <u>S</u>usceptible to infection

2. <u>I</u>nfected

3. <u>R</u>ecovered / <u>r</u>emoved (by death or quarantine).

In an SIRS model the agents then move back into a susceptible state due to new strains of the virus being introduced from an external source or due to the existing virus evolving within the network.

SIS (where the R state is ignored) and SIR household models have been used to analyze infections within sexual networks[2,3] and computer networks.[4] Typically, mean field analysis is used to determine the behaviour of the systems under varying parameters.[5–7]

1.2. Memes

The term "meme" comes from Dawkins,[8] and refers to "a unit of cultural transmission, or a unit of imitation." A meme can perhaps best be defined as "a viral idea",[9] and this is the one we use. In this paper we consider a meme to be an idea about the value of a company (as reflected in its stock price) and these memes are spread through a network. The spread of memes and their effect on the stock market has been explored by Frank, who based his memes' values (and hence efficiency of spread via imitation) on the stock price return.[10] He found that if the meme value is based on the return, then only long term fluctuations can be observed, with no explanation of shorter term fluctuations (such as those seen in "day trading"). Here we are interested in exploring these fluctuations, and hence we base our memes' values on their immediate effect on generating profits (or losses).

2. MODEL

2.1. Viruses

We model an entire school of children with grades g and classes in each grade c_g. We have several types of connections:

1. connections within a classroom, such as in Figure 1 where the squares denote students, along which there is a high probability p_1 of infection due to large amounts of time together,

2. connections between friends, both within and between classrooms, but not between students in different grades,

3. connections between a science teacher and all students of each grade (multiple classes connected together) that has separate science classes, and

4. connections between all teachers of the school.

All but the first connection are modeled with the same probability p_2 of infection due to the similar amount of time (a class or lunch time) spent together.

The type of connections are either fully connected, that is every person within a classroom connected to every other person, or connected in terms of a Moore or von Neumann neighbourhood as typically used in cellular automata. These are shown in Figure 1. When these two types of neighbourhoods are used, the students are arranged in square classes of size $x = \lfloor \sqrt{N_{c_g}} \rfloor$ where N_{c_g} is the size of class c_g. For $N \neq x^2$ we create an extra row of size $N - x^2 < x$. The probabilities used are shown in Table 1 and have largely been taken from the literature.[11, 12] Probabilities p_1 and p_2 have been estimated from the data by solving equations 1 and 2 simultaneously repeatedly over the set of times t (thus we gain an understanding of the time evolution of the probabilities). These equations are

$$p_1 \frac{I(t)}{N(t)} + p_2 \frac{I_o(t)}{N_o(t)} - p_1 \frac{I(t)}{N(t)} p_2 \frac{I_o(t)}{N_o(t)} = \frac{I(t+1) - I(t) + R(t+1)}{N(t) - R(t) - I(t)} \tag{1}$$

$$p_1 \frac{I(t+1)}{N(t+1)} + p_2 \frac{I_o(t+1)}{N_o(t+1)} - p_1 \frac{I(t+1)}{N(t+1)} p_2 \frac{I_o(t+1)}{N_o(t+1)} = \frac{I(t+2) - I(t+1) + R(t+2)}{N(t+1) - R(t+1) - I(t+1)}, \tag{2}$$

where I is the number of children infected in the class, I_o the number of children in the other classes in the grade, and similarly for N the total number of people in the class, and R is the number of children in the recovery state.

There is a high probability of infection within a class, in line with studies of transmission between children in close contact in a hospital setting.[13]

The total probability of a node being infected by virus i in a single time step is given by

$$p_{SiIi} = \min\left(1, l_{Ii}p_{1i} + f_{li}p_{2i}\right)(1 - \eta) \prod_{i \neq j} \left(1 - \left(\delta_{s_j I} + \delta_{s_j R}\right) \alpha_{ij}\right), \tag{3}$$

(a) Von Neumann neighbourhood.
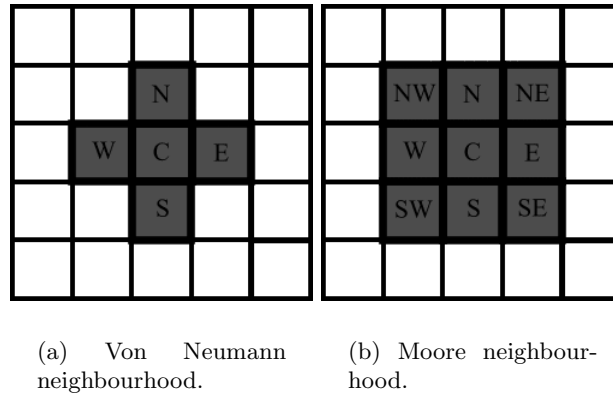
(b) Moore neighbourhood.

Figure 1. Neighbourhoods typically used in cellular automata. In the Von Neumann neighbourhood (a), every square at position (x', y') is updated based on the states of the cells in the neighbourhood $\{(x, y) \mid \text{abs}\,(x - x') + \text{abs}\,(y - y') \leq 1\}$. In the Moore neighbourhood (b), every square at position (x', y') is updated based on the states of the cells in the neighbourhood $\{(x, y) \mid \max\,(\text{abs}\,(x - x'), \text{abs}\,(y - y')) \leq 1\}$.

Table 1. Table of probabilities used for the simulation of school data. All variables have a possible range of $[0, 1]$.

Variable	Description	Default value/range explored
p_1	Probability of infection being transmitted from a neighbour in the class	x
p_2	Probability of infection being transmitted for other connections	y
p_{IR}	Probability of going from an infectious state to recovered	0.2
p_{RS}	Probability of going from the recovered state to susceptible	0.1
p_f	Probability of a person being friends with another person in the same grade	0.2
η	Individual immunity to a virus	$[0, 1)$ in steps of 0.05

where α_{ij} is the surface protein similarity between viruses i and j, and where $\delta_{s\text{X}}$ is the Kronecker delta function,

$$\delta_{s\text{X}} = \begin{cases} 1, & s = \text{X} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The surface proteins of the virus act as recognition targets for the immune system, thus having a high α means that getting one of the viruses means the immune system can (with high probability) recognise the other virus. The matrix $[\alpha_{ij}]$ is a symmetric matrix with $\alpha_{ii} = 1 \forall i$.

Levels of immunity η vary with the age of the person. We treat students within each grade g as having a fixed general level of immunity η_g. We are comparing with grades K4, K5 and 1-7. K4 and K5 are kindergarten classes containing only 4 year olds and 5 year olds, the other grades contain students of a mix of ages (maximum of a year apart in age. Values of immunity used are $\eta_{\text{K4}} = 0.0$, $\eta_{\text{K5}} = 0.05$, $\eta_{\text{teachers}} = 0.6$ and for grades 1 to 7 the immunity is given by

$$\eta_x = 0.05 + \tfrac{x}{16}, \quad x \in 1, \dots, 8, \tag{5}$$

where x is the grade, which seems to be a fit for the school data provided.

2.2. Memes

The market consists of a number of agents (traders) with the following attributes:

- Bank balance in cents (discretized in units of cents).

- Portfolio of stocks, detailing how much stock of each company the agent holds.

- A set of memes, with a maximum of one per listed company per agent.

Each meme contains the following information:

- Stock price in cents (discretized in units of cents). There is a minimum price of 1c and no maximum price. The initial stock price is set to a random number (uniform distribution) $[\max(1, S_{\text{stockname}}(0) - 100), S_{\text{stockname}}(0) + 100]$ where $S_x(t)$ is the stock price of stock x at time t.

- Volume of stocks to trade in. This is initially assigned a number in the range 0-1000 (inclusive, at random with uniform distribution).

- A counter c to keep track of how many success the meme has generated for the agent in the stock exchange. This is incremented each time an agent makes a profit and decremented each time it makes a loss. It has a minimum value of zero.

In order to transmit memes, we must define two things: a network structure, a probability of transmission of meme along links in the network, and the specific mechanism of copying. For the network structure, we considered a simple network structure where the network nodes (traders) were initially connected via directed links with probability p of a directed link from node A to node B. The memes spread only in the direction of the link. The set of p values we used for networks were $\{0, 0.5, 1\}$ for networks consisting of a single subgroup of 300 agents. These are shown in Figure 2. We also considered a scenario where the set of 300 agents was divided into two halves and one half was connected with probability $p_{11} = 0.5$, the other group with probability $p_{22} = 1.0$ (that is, fully connected) and connections from members of the first to the second group and second to the first group with probabilities $p_{12} = p_{21} = 0.2$. The copying is done by replacing an existing meme for stock x with a new meme copied along the link with probability

$$p(c) = \frac{1}{1 + \exp(c - 6)}, \tag{6}$$

where c is the counter of the success of the meme as defined above. The function is a simple sigmoid function, and is thus constrained to be in the range $[0, 1]$ as we require for a probability, and is low probability for low values of c and high for high values of c. The -6 term is there to shift the function such that it is almost zero probability for a zero counter.

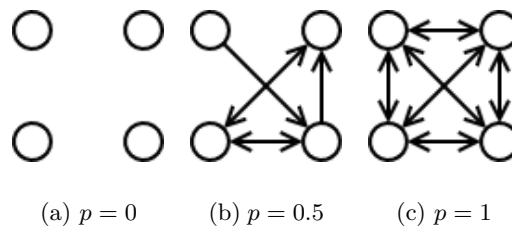(a) $p = 0$ (b) $p = 0.5$ (c) $p = 1$

Figure 2. This shows a very simple network of four stock exchange traders, with varying probabilities p of there being directed links between pairs of traders.

3. METHODS

3.1. SIRS model

To ensure the model is capturing important aspects of real infections in a school situation the model was compared with actual school data from the Colegio Nueva Granada in Bogota, Colombia, collated by the Universidad de los Andes. The data consisted of reports of infections in students from grades (years, forms) 1-7 and K4 and K5. K4 and K5 are kindergarten classes containing only 4 year olds and 5 year olds, the other grades contain students of a mix of ages (maximum of a year apart in age). Given the duration of infections, and the cold symptoms, it is most likely the infections were caused by rhinoviruses[11, 12] and not bacteria or the influenza virus.[13] The classes were approximately 20 students and there were on average 6 classes per grade. The model has therefore used a network with exactly 20 students per class and six classes. This needs to be varied in future work to better capture the actual conditions within each grade. The model directly produces information on the number of infections per person per ten week period and the numbers of people in each state in each time step.

3.2. Memes

We simulated the stock market for 100 time steps, using 300 agents. In the first simulation, each agent was given a random bank balance between 1c and $100. Since we observed inflation and this obscured the other trends (as discussed further below) we then limited the bank balance to be in the range 1c to $10. To simulate different dynamics we then used various different connection probabilities (including having two subgroups) to observe the effect of network structure and the spread of memes in different networks.

4. RESULTS

4.1. SIRS model

In Figures 3 to 6 we plot the number of people in each state over a 10 week period, from the data and our model with parameters chosen so that the plots are as close (by visual inspection) as possible. A fully connected network is used, and there is only one virus being spread through the network. As you can see, the results are similar but different in many ways. This could be due to a number of different causes such as inability to easily quantitatively match the model parameters to the real data set, the inability to assess the true network structure, or simply the stochasticity of the model.

4.2. Memes

The results with a connection probability $p = 0.5$ and bank balances set at random (uniform) in the range 1c–$100 are shown in Figure 7. You can clearly see inflation occurring. Note that although there are three shares used in all our simulation, only results for the MSFT share are shown, with initial conditions and simulation bearing no relation to the MSFT share on the NYSE. We then reduced the range to 1c–$10 and varied the connection probability from $p = 0$ (Figure 8) to $p = 0.5$ (Figure 9) and finally $p = 1$ (Figure 10). Comparing the one at $p = 0.5$ to the inflationary scenario shows how reducing the amount of free cash in the economy significantly reduces inflationary pressure, and allows other stochastic meme effects to be more visible. When the connection probability is reduced to 0, there is no spread of memes and therefore no chance of any "boom" effects. When this is increased to 0.5 a boom effect is noticeable, and finally with a connection probability $p = 1$ we are seeing some interesting dynamics with higher "boom" effects but also some significant falls in the share price. The change from one main group to two subgroups of traders further as described in Subsection 2.2 accentuates the different dynamics possible in the spread of memes, and competition between memes gives rise to the sharp boom and bust cycles shown in Figure 11. A diagram showing the spread of memes can be seen in Figure 12. This shows that the different subgroups have different dynamics. Investigation of the time-changes of the distributions of memes (not shown) reveals that one meme usually builds in popularity in one subgroup before spreading to the other subgroup, by which time the meme has already evolved in the first subgroup.

5. CONCLUSIONS AND FUTURE WORK

5.1. SIRS model

It is difficult to determine any trend at the moment. We would like to show that immunity increases with age. This seems to be a slight trend in the actual data, when matched with the current model. The current model does not match accurately enough the initial conditions of class size and number of classes per grade—the effects we are seeing could simply be due to this. Furthermore, the structure of the friendship cliques could be changing. We could try and determine this from patterns of infection within a grade, although the data is limited and this fact presents general problems for drawing strong conclusions. The data for grade K4, Figure 3(b), and seventh grade, Figure 5(b), show an interesting cycling pattern.

Both the frequency data and the SIR data suggest a similar trend in going from grade K4 to grade K5 and grade 7 to grade 8 but there is no discernible trend in going from grade K4 all the way up to grade 8. Perhaps there is a separate virus that goes through the higher grades that have immunity to the first one or the grades are just simply too "disconnected", or the data is too incomplete? We need to collect a lot more data in order to draw general statistical inferences.
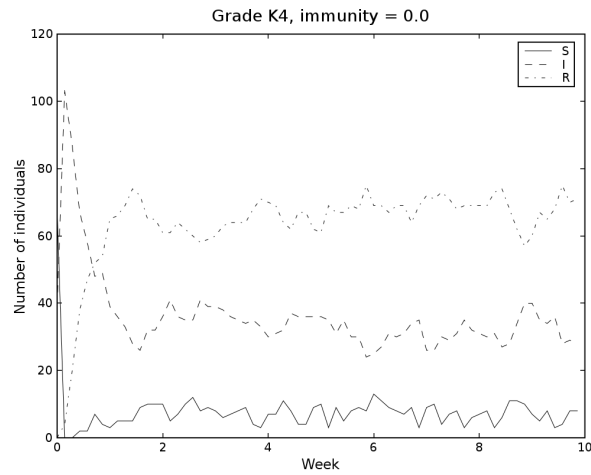
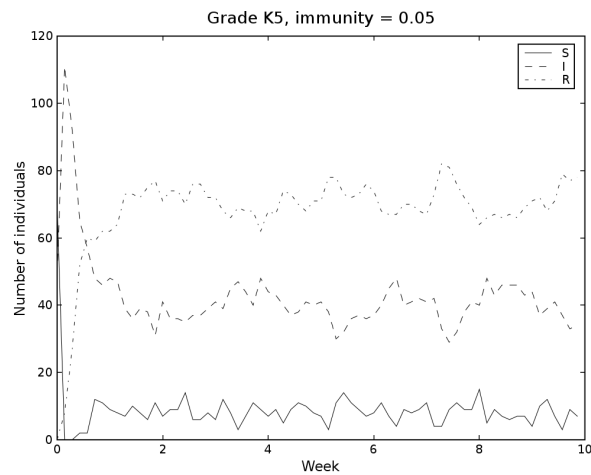(a) SIRS stages in kindergarten 4 over a 10 week period in the model.
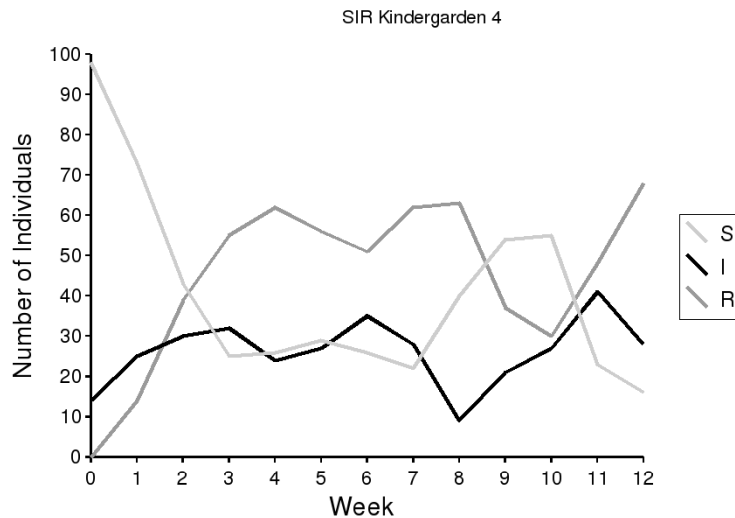
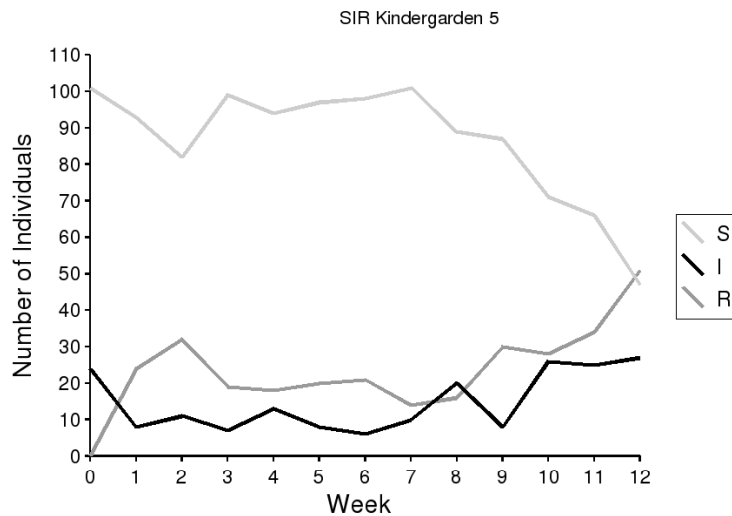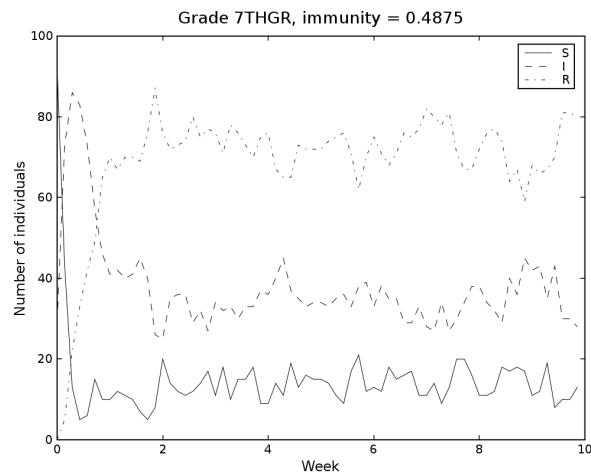(b) Actual data of SIRS stages from kindergarten 4 over a 10 week period.

Figure 3. Plot of number of people in S, I, and R stages over a 10 week period in the data and model for K4 (kindergarten, age 4) classes. Time is in days for the model, and weeks for the actual data (10 weeks = 70 days total).

Another consideration is simply better and more regular hand washing with higher and more repetition through education; education has been shown to correlate with transmission of viruses.[14] Items for future work include:

1. Include an incubation state (E).[15] Rhinoviruses have a 8-12 hour incubation period during which a person is infected but incapable of passing on the virus.[16] This is not captured in the SIRS model.

2. Make α_{ij} a function of time. One possible method is to increment α_{ij} if virus i performs worse than the

(a) SIRS stages in kindergarten 5 over a 10 week period in the model.

(b) Actual data of SIRS stages from kindergarten 5 over a 10 week period.

Figure 4. Plot of number of people in S, I, and R stages over a 10 week period in the data and model for K5 classes.

"best" virus j, as measured by total number of infections by virus j.

5.2. Memes

Clearly this model shows the effects of cash in inflation, and shows that reducing spending power is an effective way of reducing inflation. Boom effects still occur, however, as the price is governed by memes that have no relation to the underlying dynamics of business but rather due to the success of memes. Similar effects are of course seen in real life such as the recent IT stock boom where the price of the shares bore little to no relation to the underlying value, and the continual buying of stock meant more people making money, and this pattern

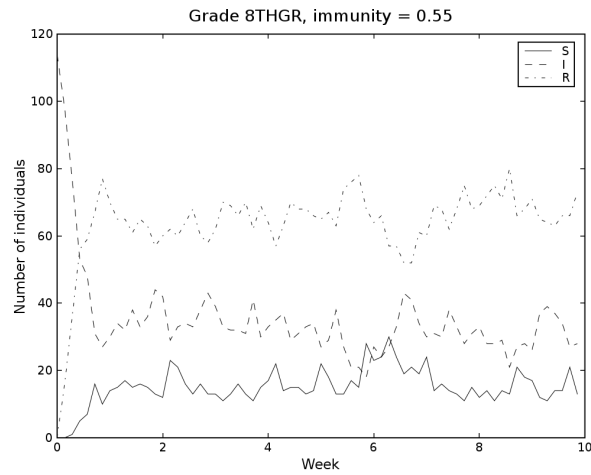(a) SIRS stages in grade 7 over a 10 week period in the model.
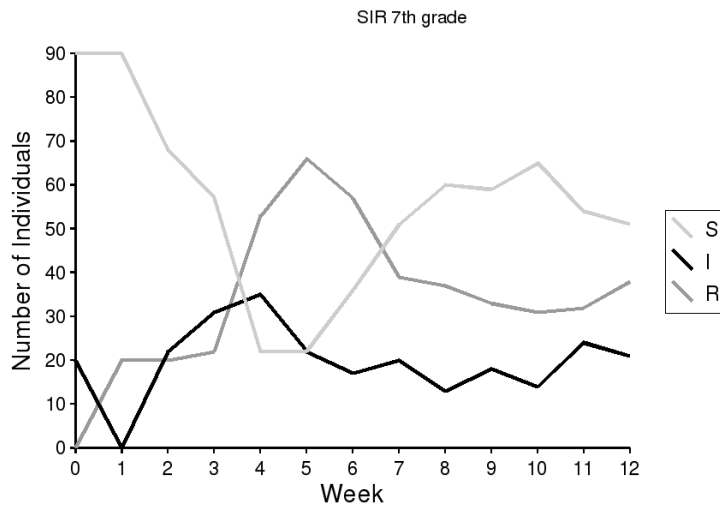
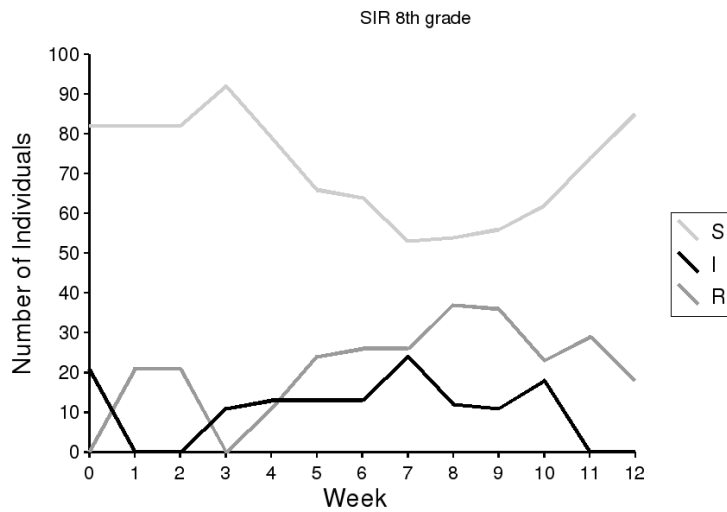(b) Actual data of SIRS stages from grade 7 over a 10 week period.

Figure 5. Plot of number of people in S, I, and R stages over a 10 week period in the data and model for grade 7 classes.

was copied until eventually it was unsustainable and prices crashed. This could be built in to the model, where the value of the meme is a function of how well the other person is doing as a whole. Other ideas for future work include:

- "Pump and dump" nodes, where one trader actively spreads a meme influencing other traders to buy, in order to sell their own stock at a higher price

- Agents that spread memes to many other traders, yet they do not trade themselves, thus mimicking newspapers and other mass media.

(a) SIRS stages in grade 8 over a 10 week period in the model.

(b) Actual data of SIRS stages from grade 8 over a 10 week period.

Figure 6. Plot of number of people in S, I, and R stages over a 10 week period in the data and model for grade 8 classes.

- Algorithms as memes, where rather than the memes being simply information, instead they contain distinct trading strategies (or mixes of multiple strategies).
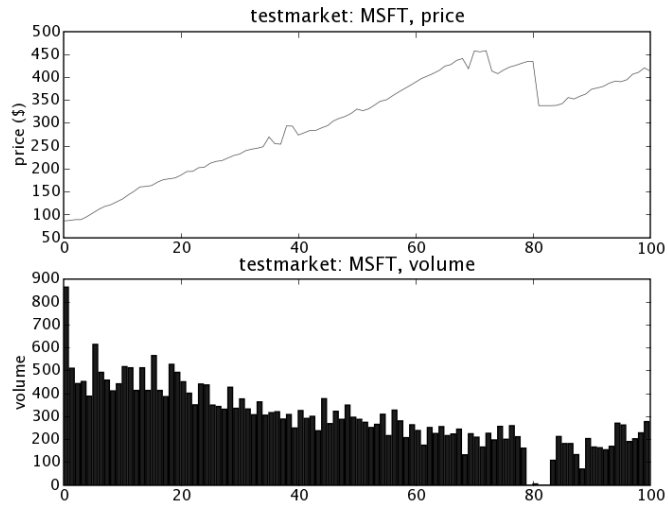
Figure 7. Graphs of the share price and volume traded when the initial bank balance is in the range 1c–$10. Observe that inflation occurs, compared with the other graphs that show either flat price trends or varying degrees of "booms" due to the spread of memes.
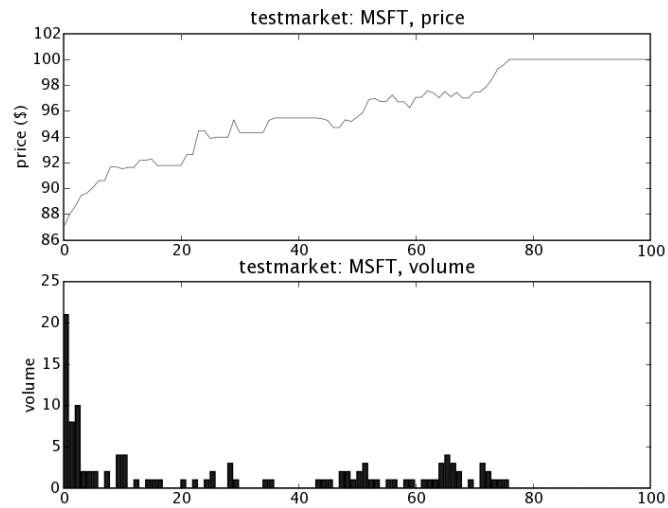

Figure 8. Graphs of the share price and volume traded when the initial bank balance is in the range 1c–$10 and the social network is totally disconnected ($p = 0$). Thus no memes spread and there are no booms or busts, merely fluctuations about the initial stock price.
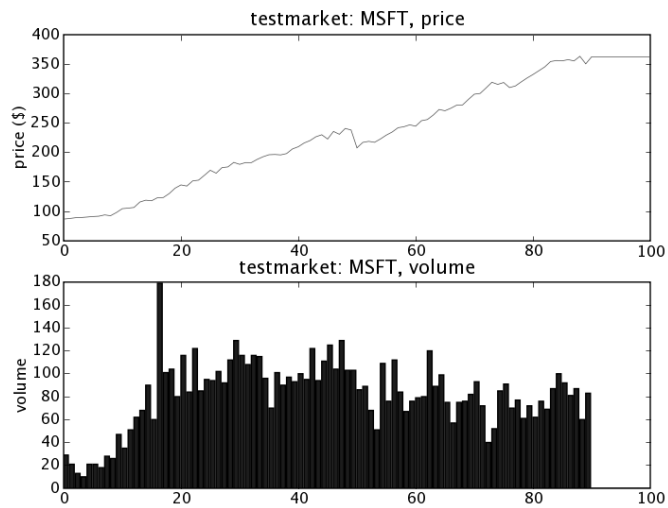
Figure 9. Graphs of the share price and volume traded when the initial bank balance is in the range 1c–$10 and the social network is totally disconnected ($p = 0.5$). Thus memes can spread and there is a boom occurring, with some fluctuations.
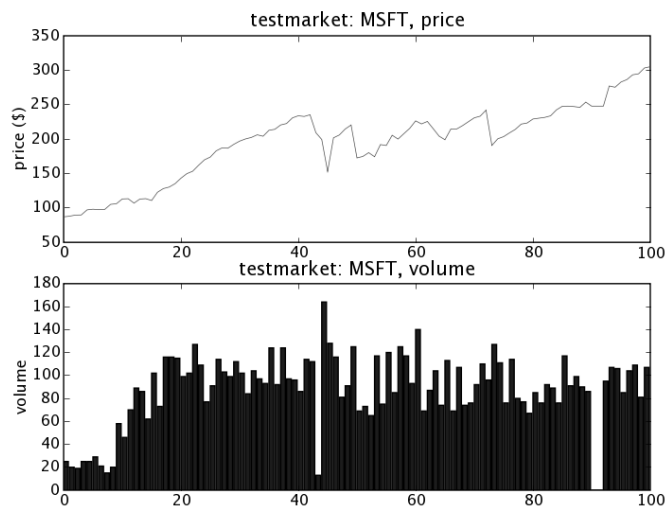

Figure 10. Graphs of the share price and volume traded when the initial bank balance is in the range 1c–$10 and the social network is fully connected ($p = 1$). The social network is large enough that differing sets of memes can start to occur and we see bust cycles in addition to boom cycles.
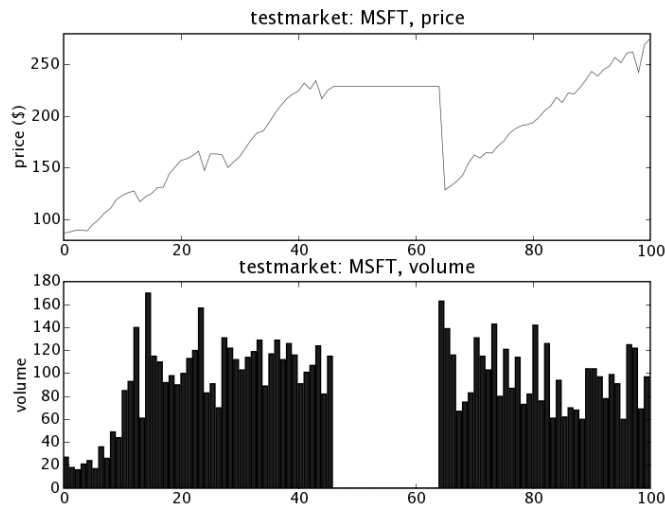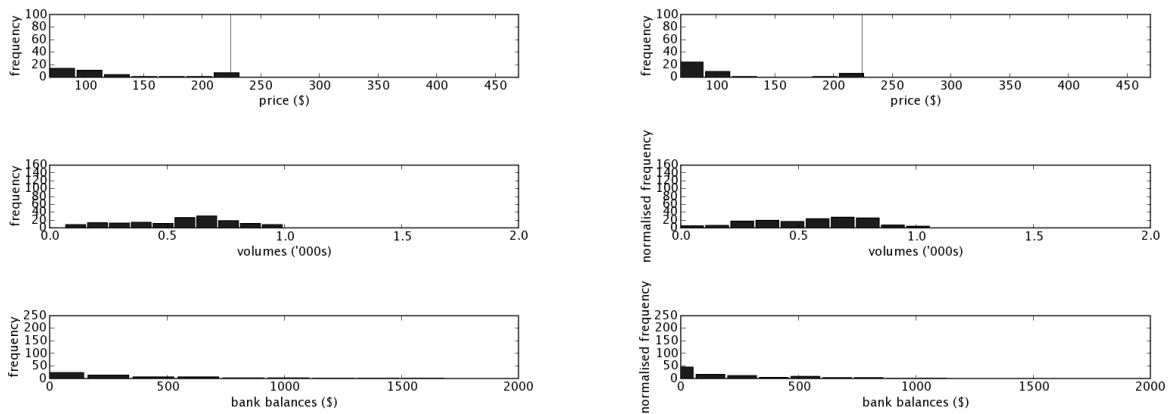
Figure 11. This shows the share price and volume traded when there are two subgroups of traders with different connection probabilities. The two subgroups (of 150 traders each) are connected with probability $p = 0.2$ of links from each group to the other. The differences in memes between the subgroups can diverge such that while one group has a buy meme, driving the stock price up, the other can initial a set of selling, driving the stock price down to the point where the other meme becomes unsustainable and then a "bust" or crash occurs. You can also see a region where the memes are two diverse, which means in the double auction market that no trades occur. Trading is still occurring in the other listed shares (graphs not shown).

(a) Histograms for subgroup with connectivity $p = 0.5$.

(b) Histograms for subgroup with connectivity $p = 1$.

Figure 12. This shows the histograms of the distribution of meme prices, meme volumes, and bank balances at time step 40 (of 100) for the two subgroups of different connection probabilities. The vertical line in the graphs of the price distributions is the current market price. The two subgroups (of 150 traders each) are connected with probability $p = 0.2$ of links from each group to the other. Here you can see there are two distinct sets of memes forming.

5.3. General conclusions and future work

The change in network structure has clear implications for the spread of both viruses and memes, and this is an area that needs further investigation. General statistical methods need to be applied to the data, with more simulation runs performed, in order to verify the effects seen and also for more quantitative comparison with real data.

6. ACKNOWLEDGMENTS

REFERENCES

1. F. G. Ball, "Threshold behaviour in stochastic epidemics among households," in *Athens Conf. Appl. Prob. and Time Series, Vol. 1: Applied Probability (Lecture Notes Statist. 114)*, C. C. Heyde, Y. V. Prohorov, R. Pyke, and S. T. Rachev, eds., pp. 253–266, 1996.
2. F. Liljeros, "Sexual networks in contemporary western societies," *Physics A* **338**, pp. 238–245, 2004.
3. J. H. Jones and M. S. Handcock, "Sexual contacts and epidemic thresholds," *Nature* **423**, pp. 605–606, June 2003.
4. J. Leveille, "Epidemic spreading in technological networks," Master's thesis, School of Cognitive and Computing Sciences, University of Sussex, UK, Sept. 2002.
5. F. Ball, "Stochastic and deterministic models for SIS epidemics among a population partitioned into households," *Mathematical Biosciences* **156**, pp. 41–67, 1999.
6. F. G. Ball, "Stochastic multitype SIR epidemics among a population partitioned into households," *Adv. Appl. Prob.* **33**, pp. 99–123, 2001.
7. G. Ghoshal and L. M. Sander, "SIS epidemics with household structure: the self-consistent field method," *Mathematical Biosciences* **190**, pp. 71–85, 2004.
8. R. Dawkins, *The Selfish Gene*, Oxford University Press, 1989.
9. N. Gaiman. Private communication, 2005.
10. J. Frank, "Applying memetics to financial markets: do markets evolve towards efficiency?," *Journal of Memetics—Evolutionary Models of Information Transmission* **3**, pp. 87–99, 1999.
11. J. M. Gwaltney Jr., *The Common Cold*, pp. 651–656. Churchill Livingstone, 5th ed. ed., 2000.
12. J. M. Gwaltney Jr., J. Hendley, G. Simon, and W. S. J. Jordon, "Rhinovirus infections in an industrial population. II. Characteristics of illness and antibody response," *JAMA* **202**, pp. 494–500, 1967.
13. D. A. Goldmann, "Respiratory infections in health-care institution," *Emerging Infectious Diseases* **7**(2), pp. 249–253, 2001.
14. Vandemoortele and E. Delamonica, "The "education" vaccine against HIV," *Current Issues in Comparative Education* **3**, pp. 1–6, Dec. 2000.
15. D. J. D. Earn, P. Rohani, B. M. Bolker, and G. B. T., "A simple model for complex dynamical transitions in epidemics," *Science* **287**, pp. 667–670, 2000.
16. J. M. Harris II and J. M. Gwaltney Jr., "Incubation periods of experimental rhinovirus infection and illness," *Clinical Infectious Diseases* **23**, pp. 1287–1290, 1996.