НАУЧНОЕ ИЗДАНИЕ МГТУ ИМ. Н. Э. БАУМАНА
НАУКА и ОБРАЗОВАНИЕ
Эл № ФС77 - 48211. Государственная регистрация №0421200025. ISSN 1994-0408
электронный научно-технический журнал

Реализация и тестирование функции хэширования данных на основе двухмерной модели Изинга

 # 02, февраль 2013
DOI: 10.7463/0213.0541576
Белим С. В., Шерешик А. Ю.

УДК 681.3

Россия, Омский государственный университет им. Ф.М. Достоевского
sbelim@mail.com

1. Введение

Большинство современных алгоритмов шифрования и хэширования основываются на ограниченном количестве хорошо известных операций, таких как подстановки, перестановки и сложение по модулю 2. Практически все блочные шифры имеют в своей основе ячейку Фейстеля. Развитие криптографических методов защиты информации идет по экстенсивному пути увеличения длины ключа и величины блока данных, что приводит к экспоненциальному росту нагрузки на аппаратную часть вычислительных систем. В связи с этим актуальной является задача поиска новых подходов к построению алгоритмов шифрования и хэширования, использующих принципиально иные подходы.

Одним из активно развивающихся в последнее время направлений является использование конечных автоматов в вопросах криптографической защиты данных. При этом главным требованием к используемым автоматам – равномерное распределение выходных данных. Таким образом, упорядоченное входное слово (открытый текст) должно преобразовываться в случайную последовательность на выходе (шифртекст). Отсюда следует, что система шифрования должна повышать энтропию преобразовываемого текста. Данным свойством обладают физические системы вследствие второго начала термодинамики. Таким образом, «хороших» свойств можно ожидать от криптографических алгоритмов, построенных на автоматах, моделирующих реальные физические системы [1, 2].

Одной из самых широко используемых моделей статистической физики является модель Изинга, позволяющая исследовать ферромагнитные материалы. Однако в

последнее время появился ряд подходов по использованию Модели Изинга в вопросах защиты информации.

Первый подход связан с использованием Модели Изинга для тестирования генераторов псевдослучайных последовательностей [3]. Вопрос получения качественной случайной последовательности весьма актуален для выработки стойких ключей шифрования, а также проверки стойкости хэш-функций и алгоритмов шифрования. Традиционный подход к тестированию генераторов псевдослучайных последовательностей связан с рассмотрением выходного набора чисел как значений случайной величины. Полученное распределение значений случайной величины анализируется статистическими тестами, основанными на некоторых предположениях. Использование же модели Изинга в качестве теста генератора псевдослучайной последовательности основано на чувствительности алгоритма Метрополиса к входной последовательности псевдослучайных чисел. Если псевдослучайная последовательность далека по своим свойствам от истинно случайной, то критические индексы системы будут отличны от предсказываемых теорией и наблюдаемых в реальном эксперименте.

Кроме тестирования случайной последовательности модель Изинга может быть использована для генерации псевдослучайной последовательности и последующего поточного шифрования на ее основе. Данная схема была реализована в работе [1] и показала достаточно хорошие результаты как по распределению выходных значений, так и по быстродействию. Ключом шифрования служили данные для инициализации последовательности.

В работе [2] для шифрования была использована модель решеточного газа, близкая к модели Изинга. В этой статье авторы приводят аналогии между этапами эволюции физической системы и этапами блочного шифрования. Также в данной работе проведен дифференциальный криптоанализ построенного алгоритма, получены оптимальные параметры для стойкости шифра.

В статье [4] модель Изинга используется для построения кода Галагера. Данный код с исправлением ошибок использует случайные матрицы, которые авторы и предлагают формировать с помощью алгоритма Метрополиса для модели Изинга. Также в работе приводится методика использования данного кода в криптосистемах с открытым ключом.

Таким образом, модель Изинга нашла применение в некоторых задачах защиты информации. Однако остается не исследованным возможность использования модели Изинга для хэширования данных. Целью данной статьи является построение и

тестирование одного из возможных вариантов алгоритма хэширования на основе двумерной модели Изинга.

2. Модель Изинга и алгоритм Метрополиса

Двумерная модель Изинга представляет собой прямоугольную сетку на плоскости, в узлах которой расположены спины S_i, принимающие одно из двух значений (1/2 или -1/2). О двух возможных значениях принято говорить как о двух противоположных ориентациях спина. Воздействие теплового движения, интенсивность которого определяется температурой, сводится к тому, что спины могут спонтанно переворачиваться в некоторые моменты времени и находиться в энергетически невыгодном положении. Общая же энергия определяется попарной суммой обменных взаимодействий между ближайшими спинами:

$$E=J \Sigma S_i S_j.$$

Здесь суммируются только пары ближайших соседей, J – константа обменного взаимодействия.

Описанная система может находиться в двух состояниях (фазах) – парамагнитном и ферромагнитном. Парамагнитная фаза соответствует неупорядоченной ориентации спинов, в результате чего суммарная намагниченность системы $m=\Sigma S_i$ будет нулевой ($m=0$). Данное состояние возможно только при наличии разупорядочивающего теплового движения. Причем тепловое разупорядочивание должно доминировать над упорядочивающим обменным взаимодействием. Парамагнитная фаза наблюдается при высоких температурах. Ферромагнитная фаза наблюдается при более низких температурах, вследствие чего намагниченность системы будет ненулевой ($|m|>0$). Температуру перехода из парамагнитной фазы в ферромагнитную (T_c) принято называть критической. Вблизи критической температуры наблюдаются критические явления, состоящие в том, что основные термодинамические функции демонстрируют сингулярное поведение, являющееся следствием неустойчивости системы.

Критическая температура перехода может быть определена с помощью кумулянтов Биндера четвертого порядка [5]:

$$U_L=1-\langle m^4\rangle/(3\langle m^2\rangle^2).$$

Для систем с разными размерами L кумулянты пересекаются в критической точке T_c.

Для моделирования поведения системы вблизи критической температуры нами был использован алгоритм Метрополиса [14]. Алгоритм Метрополиса начинается со случайно

выбранной конфигурации спинов. Затем производится случайный выбор одного из спинов и вычисляется изменение энергии ΔE. Если произошло уменьшение энергии ($\Delta E<0$), то новая конфигурация принимается. Если энергия увеличилась ($\Delta E>0$), то генерируется случайное число p из интервала [0,1] и вычисляется величина $W=exp(-\Delta E/T)$. Если $W>p$, то новая конфигурация принимается, в противном случае отбрасывается. Описанные шаги повторяются заданное количество раз. Термодинамическое усреднение производится по полученным конфигурациям.

3. Алгоритм хэширования

Рассмотрим следующий алгоритм хэширования, основанный на двумерной модели Изинга размером L×L. Двоичное представление сообщения разбиваем на блоки длиной L^2. Если длина сообщения не пропорциональна L^2, то дополняем ее случайной последовательностью. Для каждого блока проводим последовательность шагов, описанных ниже. Результирующие последовательности для каждого из блоков складываем побитово по модулю 2.

Преобразование каждого из блоков происходит по следующим шагам:

1. Используем блок в качестве начальной конфигурации спинов модели Изинга. При этом единичное значение бита соответствует спину, направленному вверх, а нулевое – спину, направленному вниз.
2. Выбираем температуру некоторую температуру системы T, которая может служить параметром хэш-функции.
3. Выполняем N шагов алгоритма Метрополиса.
4. Полученную конфигурацию спинов разворачиваем обратно в сообщение.

В данном алгоритме хэширования несколько параметров, которые необходимо выбирать опытным путем исходя из требований к стойкости хэш-функции. Как видно из алгоритма, размер хэш-значения равен L^2.

4. Компьютерный эксперимент

Построенная хэш-функция была исследована в рамках компьютерного эксперимента. Для оценки алгоритма, по результатам компьютерного эксперимента были определены значения частоты возникновения коллизий, а также лавинный эффект. Коллизией хеш-функции называют получение одинакового значения функции для разных входных данных. Коллизии существуют для большинства хеш-функций, но для качественных хеш-функций частота их возникновения близка к теоретическому

минимуму. Если хеш-функция используется для создания цифровой подписи, то умение находить для неё коллизии фактически равносильно умению подделывать цифровую подпись. Поэтому мерой криптостойкости хеш-функции считается вычислительная сложность нахождения коллизии [6]. Лавинный эффект проявляется в зависимости всех выходных битов от каждого входного бита. Обычно лавинный эффект достигается благодаря тому, что на каждом проходе изменение одного входного бита ведёт к изменениям нескольких выходных. Если криптографический алгоритм не обладает лавинным эффектом в достаточной степени, криптоаналитик может сделать предположение о входной информации, основываясь на выходной информации. Таким образом, достижение лавинного эффекта является важной целью при разработке криптографического алгоритма [7].

Были рассмотрены решетки с линейными размерами L=8 и L=16. В первом случае входящие данные были разбиты на блоки по 64 бита, во втором по 256 бит. В случае если длина входящего сообщения не была кратна длине блока, использовалась функция расширения, описанная для алгоритмов SHA-1 и MD5. Размер выходных значений также равняется 64 и 256 бит соответственно. Продолжительность моделирования для каждого блока принята равной N=1000 шагов Монте-Карло на спин. Эксперимент проводится при температуре фазового перехода T_c и более высокой температуре 1.28 T_c. В работе [8] было отмечено, что температура 1.28 T_c хорошо отражает хаотичное состояние системы, в то время как для температуры фазового перехода T_c характерна кластеризация.

В качестве тестовых данных использовали американский словарь, составленный проектом Ispell применяющийся для проверки орфографии на платформе Unix[7]. Словарь содержит 83657 слов, средняя длина которых 8,4 символа. Результаты представлены в Таблице 1.

Таблица 1. Данные о количестве обнаруженных коллизий

Алгоритм хеширования	**Размер выходного**	**Количество обнаруженных**
SHA-1	160	0
MD5	128	0
Модель Изинга (L = 8, T= T_c)	64	74796
Модель Изинга (L=16, T = T_c)	256	413
Модель Изинга (L= 8, T = 1,28T_c)	64	1986
Модель Изинга (L=16, T=1,28T_c)	256	21

Дополнительные эксперименты показали, что количество коллизий сокращается при увеличении размеров решетки. Результаты представлены на рисунке 1.

Рассмотрим насколько сильно для этих хеш-функций выражен лавинный эффект. Считается, что криптографический алгоритм удовлетворяет лавинному критерию, если при изменении одного бита входной последовательности изменяется в среднем половина выходных битов. На рисунке 1 приведена диаграмма отражающая процент изменения выходного значения при замене одного бита во входящем слове. Сокращение I(8, T_c) означает алгоритм на основе модели Изинга с линейным размером 8 и температурой T_c. Результаты получены на основании усреднения данных по 1000 пар слов отличающихся на 1 бит. Алгоритм реализован в виде компьютерной программы, обрабатывающей файлы.

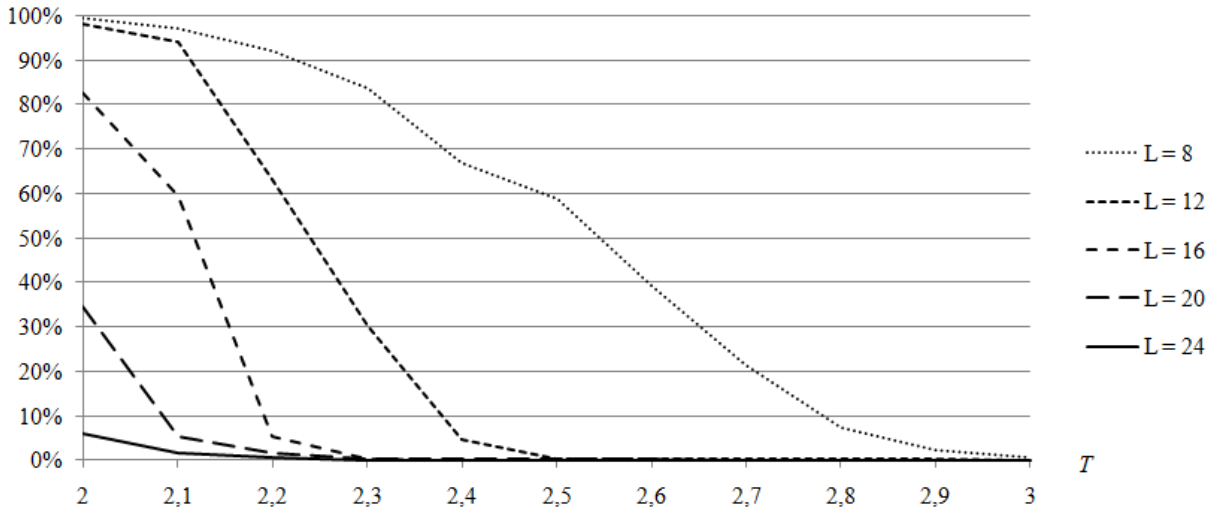

Рисунок 1. Зависимость числа коллизий от температуры и размеров решетки.

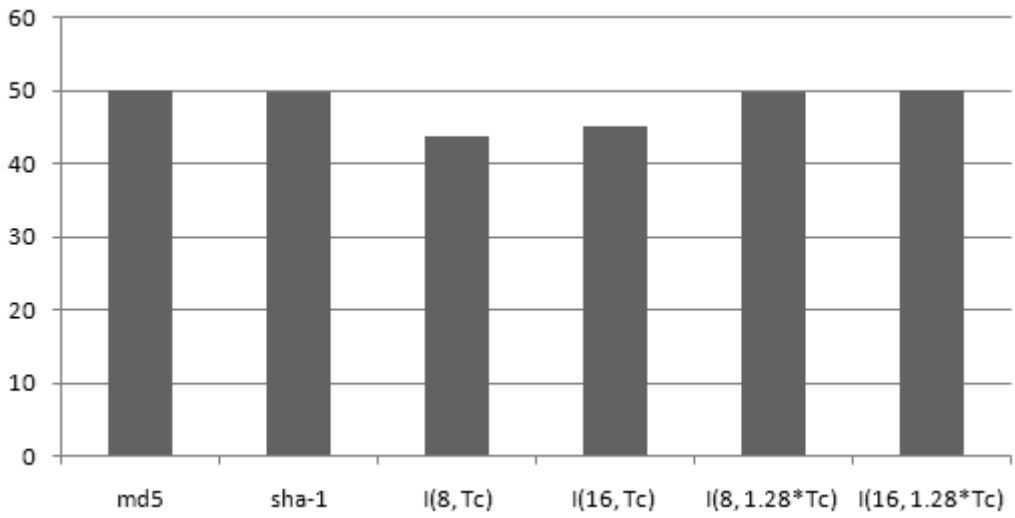

Рисунок 2. Лавинный эффект для различных функций

5. Заключение

На основании данных, полученных в результате компьютерного эксперимента, можно сделать вывод, что, при определенных условиях, алгоритм хеширования, построенный на основе модели Изинга, является достаточно качественным. Для построения эффективной хеш-функции следует использовать максимально возможный линейный размер решетки. Температура системы должна быть выше критической, чтобы система находилась в неупорядоченной фазе.

Список литературы

1. Perez A., Huynh Van Thieng C., Charbouillot S., Aziza H. Ch. 14. An En/Decryption Machine Based on Statistical Physics // Applied Cryptography and Network Security. InTech, 2012. P. 321-336. DOI: 10.5772/38808

2. Chopard B., Marconi S. Discrete Physics: a new way to look at cryptography. 2005. Available at: http://arxiv.org/pdf/nlin/0504059.pdf , accessed 01.06.2012..

3. Coddington P. D. Tests of random number generators using Ising model simulations // Int. J. Modern Physics C. 1996. Vol. 7, no. 3. P. 295-303. http://dx.doi.org/10.1142/S0129183196000235

4. Saad D., Kabashima Y., Murayama T. Public key cryptography and error correcting codes as Ising models. 2000. Available at: http://arxiv.org/pdf/cond-mat/0008363.pdf , accessed 01.06.2012.

5. Landau D.P., Binder K. A guide to Monte Carlosimulation in statistical physics. Cambridge University Press, 2005. 427 p.

6. Кнут Д. Искусство программирования. Т. 2. Получисленные алгоритмы. 3-е изд. М.:«Вильямс», 2007. 832 с.

7. Kuenning G. International Ispell Version 3.1.20. Режим доступа: http://www.cs.hmc.edu/~geoff/ispell-dictionaries.html (дата обращения 01.06.2012).

8. Fisher M.E., Barber M.N. Scaling Theory for Finite-Size Effects in the Critical Region // Phys. Rev. Lett. 1972. V. 28, no. 23. P. 1516-1519. http://dx.doi.org/10.1103/PhysRevLett.28.1516

SCIENTIFIC PERIODICAL OF THE BAUMAN MSTU
SCIENCE and EDUCATION
EL № FS77 - 48211. №0421200025. ISSN 1994-0408
electronic scientific and technical journal

Implementation and testing of hash functions based on the two-dimensional Ising model

 # 02, February 2013
DOI: 10.7463/0213.0541576
Belim S., V., Shereshik A.Yu.

Russia, Omsk F.M. Dostoevsky State University
sbelim@mail.com

The authors propose an algorithm for hashing data, basing on the increase in entropy during simulation of physical processes. The two-dimensional Ising model was selected as a system. The Ising model was studied with the use of the Metropolis algorithm. A computer experiment was carried out in order to detect collisions and determine the avalanche effect. The model's preferred parameters such as temperature and size were identified experimentally. It was shown that the avalanche effect becomes large enough when the temperature is higher than the critical one by 28%. The dimension of the system must be chosen depending on the size of the output digest. It was also shown that the two-dimensional Ising model has sufficient mixing properties for cryptographic hash functions. The authors conducted a comparison with common algorithms MD5 and SHA-1. The developed hash algorithm is scalable, unlike traditional hash algorithms.

Publications with keywords: Ising model, hash-function, algorithm to hash the data
Publications with words: Ising model, hash-function, algorithm to hash the data

References

1. Perez A., Huynh Van Thieng C., Charbouillot S., Aziza H. Ch. 14. An En/Decryption Machine Based on Statistical Physics. In book: *Applied Cryptography and Network Security*. InTech, 2012, pp. 321-336. DOI: 10.5772/38808

2. Chopard B., Marconi S. *Discrete Physics: a new way to look at cryptography*. 2005. Available at: http://arxiv.org/pdf/nlin/0504059.pdf , accessed 01.06.2012.

3. Coddington P.D. Tests of random number generators using Ising model simulations. *Int. J. Modern Physics C*, 1996, vol. 7, no. 3, pp. 295-303. http://dx.doi.org/10.1142/S0129183196000235

4. Saad D., Kabashima Y., Murayama T. *Public key cryptography and error correcting codes as Ising models*. 2000. Available at: http://arxiv.org/pdf/cond-mat/0008363.pdf , accessed 01.06.2012.

5. Landau D. P., Binder K. *A guide to Monte Carlo simulation in statistical physics*. Cambridge University Press, 2005. 427 p . http://dx.doi.org/10.1017/CBO9780511614460

6. Knuth D.E. *The Art of Computer Programming. V. 2. Seminumerical Algorithms*. Addison-Wesley, 1997. (Russ. ed.: Knut D . *Iskusstvo programmirovaniia. T. 2. Poluchislennye algoritmy*. Moscow, Vil'iams, 2007. 832 p.).

7. Kuenning G. *International Ispell Version 3.1.20*. Available at: http://www.cs.hmc.edu/~geoff/ispell-dictionaries.html , accessed 01.06.2012.

8. Fisher M.E., Barber M.N. Scaling Theory for Finite-Size Effects in the Critical Region. *Phys. Rev. Lett.*, 1972, vol. 28, no. 23, pp. 1516-1519. http://dx.doi.org/10.1103/PhysRevLett.28.1516