Plaintext Avalanche Effect
For Each Cycle of Encryption

Another Plaintext Avalanche Effect test was done, this time examining the differences between each intermediate ciphertext (the original and the bit flipped ones) at the end of each cycle of encryption. For the test, 500 original plaintexts were generated, then 128 plaintexts were generated from the originals by flipping 1 bit of the originals in each different bit position. Then, for each of 1,000 generated keys, each original plaintext was encrypted and copies of the intermediate ciphertexts of them were gathered at the end of each encryption cycle. The same was then done for each 'flipped' plaintext, where at the end of each cycle their corresponding, intermediate bits were compared with the original, intermediate ciphertext bits and the number of these differences were recorded. For each cycle, these sets of differences were output to a separate file, then a program was run on each to produce the corresponding descriptive statistics for each distribution of these differences.

Here is the hypothesis for this experiment:

The mean of the differences between the original ciphertext and the flipped ciphertext will be <= 8 bits after the 1st cycle, and then the mean of the differences will grow closer to the median of 64 bits until it is within the range of ±1 around the median by the 4th cycle, which at that point it will stay within this range for the rest of the cycles.

I am not entirely sure if this is too specific, but the claim was made due to the Diffusion Slide 2 function not being used until the 4^{th} cycle, so it seems there wouldn't be a large number of differences until then. Since there is only one bit of difference before encryption, after the first cycle the max number of differences would only be 8 bits due to there only being the 8x8 S-Box and the XOR applied the first cycle. Given that, in the middle of the 2^{nd} cycle the nybble P-Box is applied. This means that if the max difference is 8 bits, once the byte is split into nybbles and potentially moved to other positions in the intermediate ciphertext, after the XOR is applied that cycle, the max number of differences would now be 16 bits. As for the 3^{rd} cycle, it's back to just S-Box and XOR, so the max stays the same. Finally, in the middle of the 4^{th} cycle, the Diffusion Slide 2 function is applied, which effectively slides those differences across the intermediate ciphertext, which is exactly what its purpose is. Now, the hypothesis and this initial page was written before actually examining the results, so next is showing those results in the form of a large table. Afterwards, I will write up my interpretation of the results and see whether or not the hypothesis holds true.

Cycles	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Max	8	8	16	128	94	94	93	125	95	96	94	123	95	95	93	98
Min	1	1	1	1	1	1	1	1	1	2	3	10	7	8	14	32
Range	7	7	15	127	93	93	92	124	94	94	91	113	88	87	79	66
Mean	4.02	4.02	7.45	47.32	47.32	47.31	55.3	63.64	63.64	63.64	63.9	64	64	64	64	64
Mode(s)	4	4	8	48	60	60	62	64	64	64	64	64	64	64	64	64
Variance	1.94	1.95	5.62	424.5	254.2	252.9	163.4	48.55	36.52	36.45	32.96	32.12	32.01	32.02	32.01	31.99
Std. Dev.	1.39	1.39	2.37	20.6	15.94	15.9	12.78	6.97	6.04	6.04	5.74	5.67	5.66	5.66	5.66	5.66
Kurtosis	2.65	2.65	2.71	2.6	2.44	2.44	3.97	4.69	4.38	4.38	3.29	2.99	2.99	2.99	2.98	2.98
Kurt (Excess)	-0.03	-0.35	-0.29	-0.4	-0.56	-0.56	0.97	1.69	1.38	1.38	0.29	-0.01	-0.01	-0.01	-0.02	-0.02
Mid Range	4.5	4.5	8.5	64.5	47.5	47.5	47	63	48	49	48.5	66.5	51	51.5	53.5	65
MAD	1.09	1.09	1.92	16.84	13.26	13.24	9.88	5.36	4.73	4.72	4.56	4.51	4.51	4.51	4.51	4.5
RMS	4.25	4.25	7.81	51.61	49.93	49.91	56.76	64.02	63.93	63.93	64.15	64.25	64.25	64.25	64.25	64.25
Skewness	0.06	0.06	-0.19	0.26	-0.49	-0.5	-1.07	-0.32	-0.36	-0.36	-0.08	0	0	0	0	0
CV	0.35	0.35	0.32	0.44	0.34	0.34	0.23	0.11	0.1	0.09	0.09	0.09	0.09	0.09	0.09	0.09

*The count (Not shown) should be 64,000,000, but there seems to be a slight loss of precision when converting from exponential notation to integers. The loss seems negligible due to it being only a small percent of difference in comparison to the entire count, so this shouldn’t affect the quality of the results. Also, the Quartiles are always 31.5, 64, and 96.5 respectively, so the interquartile range is also always 65. The median is also always 64. Also, all decimals are rounded to 1 or 2 digits.

Results Interpretation

Restating the hypothesis for clarity:

The mean of the differences between the original ciphertext and the flipped ciphertext will be <= 8 bits after the 1st cycle, and then the mean of the differences will grow closer to the median of 64 bits until it is within the range of ±1 around the median by the 4th cycle, which at that point it will stay within this range for the rest of the cycles.

First, the mean of the differences was under 8 bits after the first cycle, and continued to be until after the 4th cycle. After that 4th cycle, however, the mean was **not** within the ±1 range of the median, and didn't fall into that range until after the 8th cycle, during which the Diffusion Slide 2 was applied for the 2nd time. Now, the mean did grow increasingly close to the median of 64 bits as it went through all the cycles. So, since the hypothesis was not correct as a whole, it seems fair to say that it is rejected.

Now, to take a closer look at the behavior of the encryption algorithm now that these descriptive statistics are gathered. The range would grow almost to maximum during every cycle that used the Diffusion Slide 2, aside from the last cycle. The mode became 64 after the 8th cycle. The excess kurtosis was negative until after the 7th cycle, then it became increasingly close to being the normal distribution until at cycle 12 it was at a negligible difference from the normal distribution (Kurtosis excess ≈ 0). By the 12th cycle the mean, median, and mode were essentially the same value of 64 bits, though the mean was off by a small decimal amount.

Conclusion

The predicted value of the mean by the 4th cycle was off, but it did reach that range after another ¼ worth of cycles. The max for the first 2 cycles was 8 bits and then was 16 bits for the 3rd cycle, as predicted. It is unclear as to why the "stretched out" range trend didn't continue with the 16th cycle where the Diffusion Slide 2 was applied again. That said, the 1 bit of change in the plaintexts eventually produces a mean difference of 64 bits in the ciphertexts. If it was a lot less or a lot more, this seems to point at the algorithm having bias, which could lead to attacks on it. I won't make any claim about this cipher being secure by any means, but having the mean and mode match the median with these sets of differences may be a desirable property for the Plaintext Avalanche Effect. Lastly, I will say it seems to be a good idea to study the more famous, modern ciphers to have something with quality to compare VIKA Zero to, as it seems I can only gather so much about my own cipher on its own, and comparisons could end up yielding more progress of this project.

stringzzz, Ghostwarez Co.
03-05-2025